

Retrieval of ice cloud properties using an optimal estimation algorithm and MODIS infrared observations. Part I: Forward model, error analysis, and information content

Chenxi Wang¹, Steven Platnick², Zhibo Zhang³, Kerry Meyer^{2,4}, and Ping Yang⁵

1. Earth System Science Interdisciplinary Center (ESSIC), University of Maryland, College Park, MD
2. NASA Goddard Space Flight Center, Greenbelt, MD
3. Department of Physics, University of Maryland - Baltimore County, Baltimore, MD
4. Goddard Earth Sciences Technology and Research (GESTAR), Universities Space Research Association, Columbia, MD
5. Department of Atmospheric Sciences, Texas A&M University, College Station, TX

For publication in the

Journal of Geophysical Research: Atmospheres

Corresponding author address: Dr. Chenxi Wang, Earth System Science Interdisciplinary Center (ESSIC), University of Maryland, College Park, MD 20740, Email: chenxi@umd.edu

Abstract

An optimal estimation (OE) retrieval method is developed to infer three ice cloud properties simultaneously: optical thickness (τ), effective radius (r_{eff}), and cloud-top height (h). This method is based on a fast radiative transfer (RT) model and infrared (IR) observations from the MODerate resolution Imaging Spectroradiometer (MODIS). This study conducts thorough error and information content analyses to understand the error propagation and performance of retrievals from various MODIS band combinations under different cloud/atmosphere states. Specifically, the algorithm takes into account four error sources: measurement uncertainty, fast RT model uncertainty, uncertainties in ancillary datasets (e.g., atmospheric state), and assumed ice crystal habit uncertainties. It is found that the ancillary and ice crystal habit error sources dominate the MODIS IR retrieval uncertainty and cannot be ignored. The information content analysis shows that, for a given ice cloud, the use of four MODIS IR observations is sufficient to retrieve the three cloud properties. However, the selection of MODIS IR bands that provide the most information and their order of importance varies with both the ice cloud properties and the ambient atmospheric and the surface states. As a result, this study suggests the inclusion of all MODIS IR bands in practice since little *a priori* information is available.

1. Introduction

Ice clouds have received great attention due to their strong impact on regional and global climate [Liou, 1986; Baran, 2012; Yang *et al.*, 2015]. To fully understand ice cloud radiative impacts and associated uncertainties, the study of cloud optical thickness (τ), effective particle radius (r_{eff}), and cloud-top height (h) on the global scale is necessary. Satellite observations provide the only means to infer global ice cloud properties. Numerous satellite instruments have been used to retrieve radiatively-relevant ice cloud properties over the past several decades [King *et al.*, 1992; Chepfer *et al.*, 1998; Minnis *et al.*, 1993, 1998, 2011; Platnick *et al.*, 2003; Huang *et al.*, 2004; Wei *et al.*, 2004; Meyer *et al.*, 2004; Heidinger and Pavolonis, 2009; Meyer and Platnick, 2010; Watts *et al.*, 2011; Baum *et al.*, 2012; Poulsen *et al.*, 2012; Walther and Heidinger, 2012; Garnier *et al.*, 2012, 2013; Kahn *et al.*, 2014]. For example, passive sensor based methods using imager and sounders include the visible and near-/shortwave-/midwave-infrared (VNIR/SWIR) bi-channel reflectance method [Nakajima and King, 1990], water vapor absorption channel cirrus retrieval method [e.g., Meyer and Platnick, 2010], the thermal infrared split-window (IR-split) method [Inoue, 1985, Parol *et al.*, 1991], and other IR methods [e.g., Minnis *et al.*, 2011; Heidinger *et al.*, 2015, Garnier *et al.*, 2012, 2013] to infer ice cloud optical and microphysical properties. The VNIR/SWIR technique is not applicable to nighttime scenes and can have higher uncertainties in high-latitude regions and optically thin cirrus cloud scenes. On the contrary, the IR-split method has two inherent advantages in the inference of cirrus clouds residing in the upper troposphere: 1) cirrus clouds reduce emission from warm surfaces and emit at a much colder temperature. As a result, radiances at the top of the atmosphere (TOA) in the presence of cirrus can significantly differ from clear-sky radiances. 2) Consistent IR observations in both daytime and nighttime allow us

to build a comprehensive cloud climatology and understand the full cloud diurnal cycle. Furthermore, cloud particle absorption strongly depends on particle size. Because of the significant ice absorption difference that occurs between the two IR window channels at $8.5\ \mu\text{m}$ and $11.0\ \mu\text{m}$, it is feasible to infer ice cloud r_{eff} from IR observations. Cloud-top height is another critical variable that determines the outgoing longwave radiation. Passive sensor cloud-top height retrieval algorithms are generally related to above-cloud gas absorption. For example, the widely employed CO_2 slicing technique uses several thermal IR bands near the $15\ \mu\text{m}$ CO_2 absorption region to infer cloud-top height [Smith and Platt, 1978; Menzel et al., 1983; Wylie and Menzel, 1999]. Since the absorptivity of CO_2 across this spectral region changes significantly, observations in these bands are sensitive to different atmospheric levels. Therefore, clouds located in different vertical regions can be retrieved with different combinations of these IR bands.

Despite this remarkable progress, ice cloud properties are still a major source of uncertainty in climate modeling [Waliser et al., 2009; Li et al., 2012]. One major reason is that cloud retrievals strongly rely on assumed cloud microphysical properties such as particle size distribution and particle habits, in particular retrievals using VNIR/SWIR observations [Zhang et al., 2009; Baum et al., 2014]. In situ measurements collected in field campaigns provide direct observational support of ice cloud microphysics and in-cloud physical processes for remote sensing studies. It is found from in situ measurements that a wide and complex variety of ice crystal sizes and habits/structures exist [Heymsfield et al., 2002; de Reus et al., 2009]. Clouds formed through different microphysical processes and under different ambient conditions will consist of ice crystals with very different size and habit distributions [Heymsfield et al., 2002; Baum et al., 2011]. In-cloud physical processes, such as melting, condensation, collision, and

84 coalescence, result in more complicated and irregular habits of ice crystals [Pfalzgraff *et al.*,
85 2010]. Complicated ice cloud microphysical properties prevent the radiative modeling of real ice
86 clouds for remote sensing applications and climate models for two reasons. First, the single-
87 scattering properties are only available for a relatively small number of geometrically simple ice
88 crystal habits (e.g., pristine hexagonal plates and columns), which cannot fully represent real ice
89 cloud particles. Second, in operational applications, only one particular habit or a certain mixture
90 of several habits is employed for global ice cloud retrievals [e.g., Platnick *et al.*, 2003; Yang *et*
91 *al.*, 2007; Minnis *et al.*, 2011; Yang *et al.*, 2013], which simplifies the retrieval algorithm but can
92 introduce a significant error source [Zhang *et al.*, 2009].

93 In addition to errors from an assumed ice crystal habit, errors from ancillary datasets
94 (e.g., atmospheric profiles, surface emission and reflectivity), satellite observations, and forward
95 radiative transfer (RT) models should be considered. However, error estimations in current ice
96 cloud retrieval methods are incomplete. Specifically, in most satellite-based ice cloud retrieval
97 algorithms, uncertainties from ice crystal microphysical properties are ignored. Although several
98 previous sensitivity studies have demonstrated that ice crystal microphysical assumptions can
99 introduce significant uncertainties to cloud retrievals [Cooper *et al.*, 2003, 2006], it is difficult to
100 estimate how much retrieval uncertainty is quantitatively contributed by an assumed habit in
101 operational applications. Furthermore, in order to mitigate the computational burden, IR-based
102 retrieval applications often estimate error statistics (e.g., covariance matrix) for ancillary errors
103 outside of the retrieval process. For example, Kahn *et al.* [2008] estimated such errors by
104 introducing Gaussian-distributed errors to atmospheric profiles and surface temperature, using
105 Atmospheric Infrared Sounder (AIRS) validation results. A large number of perturbed
106 parameters are generated, with which the perturbed forward model simulations are calculated

under different cloudy states. Statistics of simulated error due to the perturbed parameters can be evaluated by comparing the perturbed simulations against the reference. Similar methods are employed in many other retrieval studies [e.g., *Iwabuchi et al.*, 2014; *Wang et al.*, 2014]. The advantage of this method is that the generated simulation error statistics can be directly used in a retrieval algorithm and therefore computational resources are reduced. However, this method is inflexible because the statistics of the ancillary data errors are fixed. In fact, these parameter errors will have spatial and temporal variations, and strongly depend on the atmosphere and cloud conditions. For this reason, a more flexible and computationally efficient ice cloud modeling and retrieval framework is required to estimate simulation/retrieval uncertainties for various atmospheric and surface states.

In this study, we develop an optimal estimation (OE) based algorithm to simultaneously retrieve the three ice cloud properties (τ , r_{eff} , and h) using MODIS IR observations. A unique feature of this retrieval algorithm is that four different error sources are taken into account: cloud microphysical assumption errors, ancillary data errors, observational errors, and forward RT model errors. A computationally efficient forward model is employed to simulate MODIS IR observations and estimate simulation uncertainty from different error sources within the retrieval process. In this study, we wish to 1) provide a nighttime capability to complement the current MODIS (MOD06) cloud optical and microphysical property daytime-only product; and 2) improve our understanding of IR-based ice cloud retrieval uncertainties.

This paper (hereafter, Part 1) describes the forward model and retrieval algorithm, introduces the four error sources, and presents the information content analysis. A follow-up paper (hereafter Part 2) will report the retrieval results and validation against active sensor-based products. Part 1 is organized as follows: Section 2 describes the forward model, including a

clear-sky module and an ice cloud module. Section 3 introduces the error sources. Section 4 discusses the method to estimate the impact from each error source on the observational space. The OE-based retrieval algorithm is introduced in Section 5. Summary and further discussion are given in Section 6.

2. Forward Model

A computationally efficient RT model simulating the MODIS thermal IR observations is used. Only a brief description of the RT model is provided here. The present RT model is a one-dimensional, single-layered cloud model. To mitigate the computational burden, cloud scattering, emission, and absorption are considered using lookup tables (LUTs) calculated using a discrete ordinate method radiative transfer code (DISORT [Stamnes *et al.*, 1988]). Furthermore, the microphysical properties are identical throughout the whole cloud layer (homogeneous cloud assumption) but the cloud layer temperature varies linearly with height. Wang *et al.* [2011, 2013, hereafter referred to as *W11* and *W13*] discussed the IR RT equation solver of this model in detail. In the LUTs, cloud scattering properties are calculated for 16 incident angles and/or viewing angles, 40 τ , and 18 r_{eff} values (see Table 1).

In *W11* and *W13*, a correlated k -distribution code [CKD, Kratz, 1995] and a large clear-sky transmittance database are used to simulate gas absorption in MODIS IR channels and in the IR window region with a high spectral resolution, respectively. This study employs the clear-sky module of the Community Radiative Transfer Model (CRTM) designed for a wide variety of sensors including the Aqua MODIS [Han *et al.*, 2006; Liu and Weng, 2006]. By comparing the Aqua MODIS IR observations with collocated high-spectral resolution observations, such as from AIRS and IASI, some previous studies evaluated the quality of the spectral response functions (SRFs) of the Aqua MODIS IR channels and demonstrated that the Aqua MODIS

153 SRFs are shifted in some water vapor and CO₂ channels [Tobin *et al.*, 2006; Sohn *et al.*, 2010].
 154 These shifted SRFs lead to simulation biases up to 2-3 K in terms of brightness temperature
 155 (BT). The CRTM clear-sky module has corrected the shifted MODIS spectral response functions
 156 [Liu and Boukabara, 2014].

157 In this study, ice crystals are assumed to be aggregates of eight hexagonal columns
 158 with severely roughened surfaces (hereafter referred to as aggregate columns) [Yang *et al.*, 2013]
 159 and satisfy a gamma distribution with an effective variance of 0.1. This is identical to the ice
 160 particle assumption used in the Collection 6 MODIS cloud product (MOD06) [Holz *et al.*, 2015].
 161 It is necessary to emphasize that, even in the IR spectral region, ice cloud property retrievals
 162 depend on the assumed cloud microphysical properties. Holz *et al.* [2015] found that the use of
 163 aggregated columns in the IR-split window technique provides consistent cloud optical thickness
 164 retrievals in comparison with lidar and VNIR/SWIR retrievals. Other studies also looked into
 165 microphysical assumptions that provide consistent retrievals between IR, VNIR/SWIR total
 166 reflectance and polarimetric directional observations [e.g., Cole *et al.*, 2013; Baum *et al.*, 2014].
 167 Nevertheless, assuming a single habit and size distribution is not expected to represent the
 168 variety of particle shapes and distributions occurring in real ice clouds. Uncertainties for the
 169 microphysical assumption will be estimated in Section 4.

170 The forward RT model (**F**) can be expressed as a function of cloud properties and other
 171 known model input (ancillary parameters):

$$172 \quad \mathbf{Y} = [\text{BT}_1, \text{BT}_2, \dots, \text{BT}_m]^T = \mathbf{F}[\mathbf{X}(\tau, r_{\text{eff}}, h), \mathbf{P}] + \mathbf{e}, \quad (1)$$

173 where **Y** is a vector consisting of *m* MODIS IR observations in BT, **P** is a vector that includes
 174 quantities provided by ancillary datasets (e.g., air temperature (**T**), water vapor concentration

(\mathbf{w}), and ozone concentration (\mathbf{O}_3) profiles, surface emissivity spectrum ($\epsilon_{_s}$) and temperature ($T_{_s}$), and \mathbf{e} is a vector including the differences between simulation and measurement. The state vector (\mathbf{X}) includes n ($n=3$) components, namely, ice cloud optical thickness τ , effective particle radius r_{eff} , and cloud-top height h . Hereafter bold variables are vectors or matrices, unless otherwise stated. *W11* and *W13* introduced the formula and equations mapping \mathbf{X} and \mathbf{P} onto the observational space. Several Jacobian matrices are required for the retrieval and error analysis. A $m \times n$ matrix $\mathbf{K}_{\mathbf{F}/\mathbf{X}}$ consists of the partial derivatives of the m MODIS IR observations with respect to the n cloud parameters:

$$\mathbf{K}_{\mathbf{F}/\mathbf{X}} = \left[\frac{\partial \mathbf{F}}{\partial \tau}, \frac{\partial \mathbf{F}}{\partial r_{eff}}, \frac{\partial \mathbf{F}}{\partial h} \right]^T. \quad (2)$$

Figure 1 shows the output \mathbf{F} (in BT) of the present model as a function of cloud optical thickness, and the corresponding Jacobian matrix $\mathbf{K}_{\mathbf{F}/\mathbf{X}}$. With given h and r_{eff} , the TOA BTs in all MODIS IR bands decrease with an increase in τ , but with different slopes. Panels b, c, and d show the sensitivities of satellite observations in each band to the state variables. It is found that the IR measurements have the largest sensitivities to τ and r_{eff} when $0.3 < \tau < 5$ (Garnier et al., 2012), while the maximum sensitivities of BTs to h occur when $\tau > 10$. In addition, cloud optical properties in the midwave-IR band (Band 20 at $3.8 \mu\text{m}$) are different from those in the thermal IR bands. For example, BT at $3.8 \mu\text{m}$ decreases slowly with τ when $\tau < 3$, and has a maximum sensitivity to r_{eff} when $\tau > 5$. With these features, the $3.8 \mu\text{m}$ channel could provide unique ice cloud property information compared with thermal IR observations (see Section 5). Simulation of daytime observations in this band is complicated due to a non-trivial solar reflectance signal, which is not included in the present forward RTM.

Other Jacobian matrices include the first derivative of BTs with respect to ancillary data parameters. For example, $\mathbf{K}_{\mathbf{F}/\mathbf{I}_s}$ indicates the derivative of BTs with respect to surface emission spectrum (\mathbf{I}_s); $\mathbf{K}_{\mathbf{F}/\mathbf{B}(\mathbf{T})}$ and $\mathbf{K}_{\mathbf{F}/\tau_g}$ are Jacobian matrices of air emission and optical thickness (τ_g), respectively. The matrix $\mathbf{K}_{\mathbf{F}/\mathbf{B}(\mathbf{T})}$ of an l -layer atmosphere can be expressed as:

$$\mathbf{K}_{\mathbf{F}/\mathbf{B}(\mathbf{T})} = \left[\frac{\partial \mathbf{F}}{\partial \mathbf{B}(T_1)} \quad \frac{\partial \mathbf{F}}{\partial \mathbf{B}(T_2)} \quad \cdots \quad \frac{\partial \mathbf{F}}{\partial \mathbf{B}(T_l)} \right]^T, \quad (3)$$

where \mathbf{B} is the Planck function, T_i indicates the temperature of the i^{th} atmospheric layer.

Derivations of Jacobian matrices $\mathbf{K}_{\mathbf{F}/\mathbf{X}}$, $\mathbf{K}_{\mathbf{F}/\mathbf{I}_s}$, $\mathbf{K}_{\mathbf{F}/\mathbf{B}(\mathbf{T})}$, and $\mathbf{K}_{\mathbf{F}/\tau_g}$ are given in Appendices A and B.

3. Optimal Estimation based retrieval method

The optimal estimation (OE)-based retrieval method [Rodgers, 2000] is an efficient inversion method widely applied to a wide variety of remote sensing applications [Poulsen *et al.*, 2012; Sourdeval *et al.*, 2013, 2015; Iwabuchi *et al.*, 2014]. The retrieval is essentially a process that reduces the state vector uncertainty from the *a priori* state, which depends on our knowledge of the state variables, to the *posterior* state when measurements are made. The OE method retrieves the state variables having the maximum probability of occurrence by minimizing a cost function J :

$$J = [\mathbf{F}(\mathbf{X}, \mathbf{P}) - \mathbf{Y}]^T \mathbf{S}_y^{-1} [\mathbf{F}(\mathbf{X}, \mathbf{P}) - \mathbf{Y}] + [\mathbf{X} - \mathbf{X}_a]^T \mathbf{S}_a^{-1} [\mathbf{X} - \mathbf{X}_a], \quad (4)$$

where \mathbf{X}_a and \mathbf{X} are the *a priori* and *posterior* state vectors, \mathbf{S}_y and \mathbf{S}_a are covariance matrices of the observation to simulation differences and the uncertainty of the *a priori* state vector, respectively. Generally, we assume large *a priori* uncertainties so that the cost function J is

dominated by the first term of equation 4. Minimizing the cost function J is a non-linear least squares fitting problem. The Levenberg-Marquardt iteration method [Levenberg, 1944; Marquardt, 1963] is an efficient approach to solve this problem and is used in this study. This method is a combination of the gradient descent method and the Gauss-Newton method, expressed as follows:

$$\Delta \mathbf{X}_i = \left(\mathbf{K}_{\mathbf{F}/\mathbf{X}_i}^T \mathbf{S}_{y,i}^{-1} \mathbf{K}_{\mathbf{F}/\mathbf{X}_i} + \gamma_i \mathbf{I} \right)^{-1} \mathbf{K}_{\mathbf{F}/\mathbf{X}_i} [\mathbf{Y} - \mathbf{F}(\mathbf{X}_i)], \quad (5)$$

where i indicates the i^{th} iteration, γ is a positive damping parameter and varies at each iteration according to the variation of the cost function. Generally speaking, in this retrieval algorithm, γ varies from 0.01 and is decreased (or increased) by a factor of 5 if a new cost function is smaller (or larger) than previous iteration step. As γ approaches 0, Eq. (5) essentially reduces to the Gauss-Newton method, while it approaches the gradient descent method for large γ . Marquardt [1963] found that since elements in the state vector \mathbf{X} could have different magnitudes, the identity matrix \mathbf{I} in Eq. (5) could be replaced by a scaling matrix $\mathbf{D} = \mathbf{S}_a^{-1}$. The retrieval stops at the p^{th} iteration (\mathbf{X}_p) when the iteration converges or the simulation fits the measurement [Rodgers, 2000]:

$$(\mathbf{X}_p - \mathbf{X}_{p-1})^T \mathbf{S}_p^{-1} (\mathbf{X}_p - \mathbf{X}_{p-1}) \ll n, \quad (6)$$

$$[\mathbf{F}(\mathbf{X}_p) - \mathbf{Y}]^T \mathbf{S}_y^{-1} [\mathbf{F}(\mathbf{X}_p) - \mathbf{Y}] \approx m, \quad (7)$$

where n and m are the number of state variables and measurements, and \mathbf{S}_p is the error covariance matrix of the *posterior* state vector \mathbf{X}_p , defined in the form:

$$\mathbf{S}_p = \left(\mathbf{S}_a^{-1} + \mathbf{K}_{F/X}^T \mathbf{S}_y^{-1} \mathbf{K}_{F/X} \right)^{-1}. \quad (8)$$

4. Error analysis

Four error sources are taken into account in the retrieval algorithm: measurement uncertainty, fast RT model uncertainty, uncertainties in ancillary datasets, and uncertainties associated with cloud particle property assumptions.

4.1 Measurement Uncertainty

Measurement uncertainty and the corresponding long-term trend of MODIS due to the instrumental noise and degradation are documented for the latest MODIS Collection 6 L1b product. Specifically, the noise-equivalent temperature differences in Aqua MODIS IR bands are generally less than 0.3 K [Xiong *et al.*, 2009]. Scaled uncertainty indices are provided in the product for individual bands and pixels. Shifted SRFs of Aqua MODIS have been found in Bands 27, 28, 34, 35, and 36, resulting in observational biases. The impact of the shifted SRFs has been evaluated quantitatively in previous studies [e.g., Tobin *et al.*, 2006; Sohn *et al.*, 2010], and can be removed by using the latest CRTM clear-sky module designed specifically for Aqua MODIS [Liu and Boukabara, 2014]. It is assumed that observational uncertainties in different bands are independent. Therefore, in this study, a $m \times m$ diagonal matrix (\mathbf{S}_{obs}) is used for each pixel to indicate the error covariance of measurement error:

$$\mathbf{S}_{obs} = \begin{bmatrix} \sigma_{obs,1}^2 & 0 & \cdots & 0 \\ 0 & \sigma_{obs,2}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{obs,m}^2 \end{bmatrix}, \quad (9)$$

where $\sigma_{obs,i}$ indicates the standard deviation of the measurement errors in band i .

4.2 Fast RT Model Uncertainty

The fast RT model uncertainty is the difference between the employed fast model, based on a set of LUTs to describe cloud bulk scattering properties, and 1-D RTMs that rigorously solve RT equations in plane-parallel atmospheres (e.g., the DISORT). It is important to emphasize that, even if perfect measurements are made and there is no input error, simulations of the rigorous 1-D RTMs can be different from observations. For example, cloud inhomogeneity effects can introduce significant retrieval errors [e.g., *Iwabuchi and Hayasaka*, 2002; *Marshak et al.*, 2006; *Fauchez et al.*, 2014, 2015; *Cornet et al.*, 2005]. However, estimation of these errors in practice is difficult and beyond the scope of this paper. In this study, a large number of forward simulations from the present RT model and DISORT under different cloud and atmospheric conditions, and viewing geometries are used to create the statistics of the fast RT model error, which are described using an $m \times m$ error covariance matrix \mathbf{S}_{RT} .

4.3 Uncertainties in Ancillary Datasets

Errors associated with non-retrieved variables dominate the IR based retrieval uncertainty and cannot be ignored [*Cooper et al.*, 2006; *Iwabuchi et al.*, 2014]. In this study, an ancillary parameter error covariance matrix (\mathbf{S}_{anc}) is used to estimate these errors quantitatively for each cloudy pixel. Specifically, uncertainties of four ancillary parameters are considered: the surface temperature (T_s), the surface emissivity (ϵ_s), and the atmospheric profiles of temperature (\mathbf{T}) and water vapor concentration (\mathbf{w}).

The surface emission (\mathbf{I}_{-s}) is the product of $\mathbf{B}(T_{-s})$ and ϵ_{-s} . The covariance matrix (\mathbf{COV}) of surface emission uncertainty ($\delta\mathbf{I}_{-s}$), under the assumption that the T_{-s} and ϵ_{-s} are independent, can be expressed as:

$$\mathbf{COV}(\delta\mathbf{I}_{-s}) = \mathbf{K}_{\mathbf{I}_{-s}/\mathbf{B}(T_{-s})} \mathbf{COV}[\delta\mathbf{B}(T_{-s})] \mathbf{K}_{\mathbf{I}_{-s}/\mathbf{B}(T_{-s})}^T + \mathbf{K}_{\mathbf{I}_{-s}/\epsilon_{-s}} \mathbf{COV}(\delta\epsilon_{-s}) \mathbf{K}_{\mathbf{I}_{-s}/\epsilon_{-s}}^T, \quad (10)$$

where $\delta\mathbf{B}(T_{-s})$ and $\delta\epsilon_{-s}$ are uncertainties of black surface emission and emissivity, respectively, $\mathbf{K}_{\mathbf{I}_{-s}/\mathbf{B}(T_{-s})}$ and $\mathbf{K}_{\mathbf{I}_{-s}/\epsilon_{-s}}$ are Jacobian matrices of surface emission with respect to $\mathbf{B}(T_{-s})$ and ϵ_{-s} , respectively. The ancillary parameter uncertainty covariance matrix due to surface emission (\mathbf{S}_{anc_s}) is:

$$\mathbf{S}_{anc_s} = \mathbf{K}_{\mathbf{F}/\mathbf{I}_{-s}} \mathbf{COV}(\delta\mathbf{I}_{-s}) \mathbf{K}_{\mathbf{F}/\mathbf{I}_{-s}}^T, \quad (11)$$

where $\mathbf{K}_{\mathbf{F}/\mathbf{I}_{-s}}$ is the Jacobian matrix of TOA BT with respect to surface emission.

For simplification, we only consider the uncertainties of air temperature and water vapor concentration in the cloud-free part of the profile, although CO_2 , O_3 , and other trace gas concentration uncertainties impact simulations in several MODIS bands, in particular MODIS bands 33-36 ($13.3\text{-}14.2\ \mu\text{m}$). Meanwhile, the relation between \mathbf{T} and \mathbf{w} profile uncertainties ($\delta\mathbf{T}$ and $\delta\mathbf{w}$) is not well quantified. In this study, we simply assume that the two uncertainties are independent. Gas emission is determined by blackbody emission $\mathbf{B}(\mathbf{T})$ and gas optical thickness τ_g . Here we separately consider the impacts from $\delta\mathbf{B}(\mathbf{T})$ and $\delta\tau_g$.

Both $\delta\mathbf{T}$ and $\delta\mathbf{w}$ profiles result in τ_g uncertainties. With independent $\delta\mathbf{T}$ and $\delta\mathbf{w}$ profiles, \mathbf{COV} of $\delta\tau_g$ can be expressed as:

$$\mathbf{COV}(\delta\tau_g) = \mathbf{K}_{\tau_g/T} \mathbf{COV}(\delta\mathbf{T}) \mathbf{K}_{\tau_g/T}^T + \mathbf{K}_{\tau_g/w} \mathbf{COV}(\delta\mathbf{w}) \mathbf{K}_{\tau_g/w}^T, \quad (12)$$

where $\mathbf{K}_{\tau_g/T}$ and $\mathbf{K}_{\tau_g/w}$ are the Jacobian matrices of τ_g with respect to \mathbf{T} and \mathbf{w} , respectively. It is important to emphasize that $\mathbf{COV}(\delta\mathbf{T})$ and $\mathbf{COV}(\delta\mathbf{w})$ are not diagonal matrices since both $\delta\mathbf{T}$ and $\delta\mathbf{w}$ have vertical correlations. Similarly, \mathbf{COV} of clear-sky blackbody emission uncertainty $\delta B(T)$ is:

$$\mathbf{COV}[\delta\mathbf{B}(\mathbf{T})] = \mathbf{K}_{\mathbf{B}(\mathbf{T})/T} \mathbf{COV}(\delta T) \mathbf{K}_{\mathbf{B}(\mathbf{T})/T}^T, \quad (13)$$

where $\mathbf{K}_{\mathbf{B}(\mathbf{T})/T}$ is the Jacobian matrix of the Planck function with respect to temperature. The ancillary parameter uncertainty covariance matrix due to $\delta\mathbf{B}(\mathbf{T})$ and $\delta\tau_g$ (\mathbf{S}_{anc_B} and $\mathbf{S}_{anc_T_g}$) can be expressed as:

$$\mathbf{S}_{anc_T_g} = \mathbf{K}_{F/\tau_g} \mathbf{COV}(\delta\tau_g) \mathbf{K}_{F/\tau_g}^T, \quad (14)$$

$$\mathbf{S}_{anc_B} = \mathbf{K}_{F/B(T)} \mathbf{COV}[\delta\mathbf{B}(\mathbf{T})] \mathbf{K}_{F/B(T)}^T, \quad (15)$$

where \mathbf{K}_{F/τ_g} and $\mathbf{K}_{F/B(T)}$ are two Jacobian matrices of TOA BT with respect to $B(T)$ and τ_g .

Assuming independence of all error components, the total ancillary parameter uncertainty can be expressed as the summation of the three components, namely the surface uncertainty, the gas optical thickness uncertainty, and air temperature (or blackbody emission) uncertainty:

$$\mathbf{S}_{anc} = \mathbf{S}_{anc_s} + \mathbf{S}_{anc_T_g} + \mathbf{S}_{anc_B}. \quad (16)$$

Derivations of Jacobian matrices $\mathbf{K}_{\tau_g/T}$, $\mathbf{K}_{\tau_g/w}$, \mathbf{K}_{F/τ_g} , and $\mathbf{K}_{F/B(T)}$ in Eqs. (11)–(15) are given in Appendix B. Figures 2 and 3 show the three components of \mathbf{S}_{anc} for two cases. In the first case,

we use a 0.5 K uncertainty of T_s and a 0.01 uncertainty of ϵ_s for all IR bands. The air T and w uncertainties are 1 K and 15 %, respectively, for all layers. In the second case, the surface temperature and emissivity uncertainties are 2 K and 0.03, respectively. The air T and w uncertainties are 2 K and 25 %, respectively, for all layers. For simplification, we assume that the ϵ_s uncertainties in all bands are independent. In both cases, the correlation between the T (or w) uncertainties of two vertical layers decreases exponentially with the increase of layer vertical distances and approaches 0 if the distance exceeds 10 km. In the window bands (e.g., Bands 20 (3.75 μm), 29 (8.6 μm), 31 (11 μm), and 32 (12 μm)) uncertainties resulting from surface temperatures and emissivity spectra are important. In absorption bands, uncertainties of air temperature and water vapor profiles lead to larger uncertainties in the gas optical thickness profile, which then result in larger simulation uncertainties. This is because water vapor is a major absorptive species, and both the absorption and emission abilities of water vapor are affected by air temperature. As a result, in these bands, simulation uncertainties are highly correlated. The ice cloud properties used to calculate these covariance matrices are: $\tau = 1$, $r_{eff} = 20 \mu\text{m}$, and $h = 10 \text{ km}$.

4.4 Ice Cloud Particle Habit Uncertainty

Uncertainties in ice crystal habits introduce additional errors to ice cloud retrievals. However, it is difficult to estimate this uncertainty directly since particle habit is not an explicit physical quantity in the radiative transfer model. In our forward RTM, the change of ice particle habit is equivalent to the change of cloud layer scattering properties, such as emissivity, transmissivity, and reflectivity, which have further impact on simulations and retrievals of ice clouds. Therefore, discussions in this section are essentially focused on uncertainties of cloud

layer scattering properties considering a large cloud particle habit ensemble. We intend to establish a general estimation of simulation uncertainty using current knowledge of in-situ observed ice crystal habits and size distributions, and theoretically derived single scattering properties of ice particles with simple habits. Obviously, the limited ice particle habits used in theoretical studies cannot fully represent the irregular ice particle habits in real ice clouds. However, theoretical habit models can nevertheless shed light on the range of cloud scattering property uncertainties.

The *Yang et al.* [2013] ice crystal single-scattering database used here includes 9 non-spherical ice crystal shapes and 3 degrees of particle surface roughness. The database covers wide ranges of ice particle maximum dimension (between 2 and 10,000 μm) and wavelength (between 0.2 and 100 μm). Additionally, the effect of particle surface roughness is simulated by randomly distorted surface slope for each incident ray in the IGOM approach. The slopes of a tilted surface facet along two orthogonal directions are specified in terms of the two-dimensional Gaussian distribution [*Yang and Liou*, 1998; *Liu et al.*, 2013] following *Cox and Munk* [1954], with a roughness parameter σ (i.e., the standard deviation of the Gaussian distribution). Specifically, three degrees of surface roughness are considered, namely smooth ($\sigma = 0$), moderately roughened ($\sigma = 0.03$), and severely roughened ($\sigma = 0.5$).

Assuming that particles with the same shapes but different degrees of surface roughness can be considered different habits, there are 27 habits in the database. With these existing habits, we can estimate how cloud scattering property uncertainties impact RT simulations with a cloud scattering property matrix (**C**). Specifically, **C** is a $mN_p \times N_h$ matrix, where m is the number of measurements, N_p is total the number of cloud scattering properties

355 (e.g., cloud layer emissivity, transmissivity, and reflectivity calculated for 16 angles, 40 τ , and 18
 356 r_{eff} values, as shown in Table 1) used in the RT simulation, and N_h is the number of possible ice
 357 crystal habits. As discussed in previous studies, the impact of an ice cloud layer in the forward
 358 RTM is primarily controlled by cloud layer emissivity (ε_c), transmissivity (Γ_c), and
 359 reflectivity (R_c). For given cloud τ and r_{eff} values, these bulk properties are pre-computed at
 360 different angles. Therefore, dozens of cloud scattering properties are involved in the cloudy-sky
 361 RT simulation. The matrix \mathbf{C} can be expressed as follows:

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} & \cdots & \mathbf{C}_{1N_h} \\ \mathbf{C}_{21} & \mathbf{C}_{22} & \cdots & \mathbf{C}_{2N_h} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}_{m1} & \mathbf{C}_{m2} & \cdots & \mathbf{C}_{mN_h} \end{bmatrix}, \quad (17)$$

363 where \mathbf{C}_{ij} is a vector that includes the cloud scattering properties of the j^{th} particle habit in the i^{th}
 364 band:

$$\mathbf{C}_{ij} = [c_{ij}^1, c_{ij}^2, \dots, c_{ij}^{N_p}]^T, \quad (18)$$

366 where c indicates the cloud scattering property, and the superscript denotes the index of a cloud
 367 scattering property parameter.

368 The uncertainty covariance matrix of the assumed ice crystal habits (\mathbf{S}_{habit}) is given by

$$\mathbf{S}_{habit} = \mathbf{K}_{F/C} \mathbf{COV}(\mathbf{C}) \mathbf{K}_{F/C}^T, \quad (19)$$

370 where the $m \times mN_p$ matrix $\mathbf{K}_{F/C}$ is the Jacobian matrix of the simulation with respect to cloud
 371 scattering properties. Derivation of $\mathbf{K}_{F/C}$ is given in Appendix C. The covariance matrix $\mathbf{COV}(C)$
 372 is a $mN_p \times mN_p$ matrix.

373 The covariance matrix (\mathbf{S}_y) of the total measurement to simulation difference is the
 374 summation of the four components if the four types of error are independent:

$$375 \quad \mathbf{S}_y = \mathbf{S}_{obs} + \mathbf{S}_{RT} + \mathbf{S}_{anc} + \mathbf{S}_{habit} . \quad (20)$$

376 Figure 4 shows these components of \mathbf{S}_y except \mathbf{S}_{RT} . In comparison with the other three
 377 components, \mathbf{S}_{RT} is too small to be visualized. The ice cloud state is the same as the one used in
 378 Figures 2 and 3. \mathbf{S}_{anc} is identical to that shown in Figure 2, and \mathbf{S}_{obs} reflects a typical magnitude
 379 of MODIS observational uncertainty. Figure 5 shows the fractional contributions of the 4
 380 different errors to the diagonal entries of matrix \mathbf{S}_y . It shows that the maximum impact of
 381 ancillary parameter errors occurs in absorption bands, in particular Band 27 and 36 (yellow bars
 382 in Figure 5). In window bands, such as Bands 20, 29, 31, and 32, although the contribution from
 383 \mathbf{S}_{anc} is also important, uncertainty due to ice crystal habit assumption (\mathbf{S}_{habit}) is comparable to
 384 \mathbf{S}_{anc} . Contributions from \mathbf{S}_{obs} and \mathbf{S}_{RT} are small in magnitude, in comparison with \mathbf{S}_{anc} and \mathbf{S}_{habit} ,
 385 suggesting that the satellite measurement error and the fast model error are small. Uncertainties
 386 from the four different sources will influence retrieval accuracy, and are shown in Figure 6. As
 387 expected, uncertainties from ancillary datasets and ice crystal habits are two important
 388 contributors and should not be ignored.

389 5. Information content analysis

The Shannon information content (hereafter information content), defined by *Shannon and Weaver* [1949], is used to describe the entropy reduction of a variable (or a set of variables) after measurement. *Rodgers* [2000] showed that the entropy (S) of a multivariate Gaussian distribution with covariance matrix \mathbf{S} is

$$S = \frac{1}{2} \ln |\mathbf{S}| + \text{constant}. \quad (21)$$

The information content (H) is defined as

$$H = \frac{1}{2} \ln |\mathbf{S}_a| - \frac{1}{2} \ln |\mathbf{S}_p| = \frac{1}{2} \ln |\mathbf{S}_a \mathbf{S}_p^{-1}|, \quad (22)$$

where \mathbf{S}_a and \mathbf{S}_p are the error covariance matrices of the prior and posterior state vector \mathbf{X} as discussed in Section 4. Substituting Eq. (8) in Eq. (22) we have

$$H = \frac{1}{2} \ln \left| \mathbf{S}_a \left(\mathbf{S}_a^{-1} + \mathbf{K}_{\mathbf{F}/\mathbf{X}}^T \mathbf{S}_y^{-1} \mathbf{K}_{\mathbf{F}/\mathbf{X}} \right) \right|. \quad (23)$$

A binary logarithm is frequently used in information theory when the unit of information is “bit”. The information content H of a measurement quantitatively indicates the factor by which the uncertainty of retrieval variables decreases with the measurement. From Eq. (22), it is obvious that H is 0 if a useless measurement is made, which does not decrease the uncertainty of the a priori estimation. A positive H indicates the knowledge of retrieval variables benefits from a measurement. Eq. (23) shows that the information content not only depends on the sensitivity of the measurement to the retrieval variables ($\mathbf{K}_{\mathbf{F}/\mathbf{X}}$), but also on the errors of the forward model and measurements (\mathbf{S}_y), and the a priori estimation (\mathbf{S}_a). The framework of information content has

been frequently used to evaluate the sensitivity of a remote sensing method or algorithm to model parameters [e.g., L’Ecuyer, et al. 2006, Cooper et al. 2006; Coddington et al., 2012].

Figures 7 and 8 show information content analyses associated with different ice cloud states. In each panel of the figures, the H of a single band is calculated first, as indicated by the black curve. Large H indicates better estimation of the three retrieval variables. With the help of one MODIS band, the performance of each band varies with cloud properties. For example, in comparison with other bands, Band 31 has the largest H for optically thin cirrus cloud consisting of relatively small particles because BT in Band 31 is sensitive to both τ and r_{eff} (Panels a and d of Fig. 7); whereas H for Band 20 improves with the increase of τ because the cloud is more “transparent” in Band 20 (Panels c and f of Fig. 7). For each panel, we first find the best measurement (band) from the black curve. Based on that first band, we continually calculate H values for two measurements for each ice cloud state (see blue curve). The blue curve indicates the information content of the second band and the previously determined first band. Consequently, the second best band can be found from the blue curve. Similar processes can be done to find the third, forth, and fifth best bands (illustrated with green, yellow, and red curves, respectively). Obviously, the overall information content increases with an increase in the number of measurements. However, H increases with the number of measurements slowly when more than four bands are involved, suggesting that the MODIS IR bands are not independent. Furthermore, it is found that under some cloud states, the overall H values are relatively low even if the best five bands are selected. For example, red curves in the first columns of Figures 7 and 8 show the information contents obtained from the best five bands for optically thin cirrus clouds with $\tau = 0.1$. The overall H values are approximately 3, suggesting that the largest IR retrieval uncertainty occurs when the cloud is optically thin.

Figure 9 shows the change in retrieval uncertainty with an increase in the number of measurements. Generally, retrieval uncertainty decreases and approaches the minimum if more than three channels are used. Here the uncertainty of $\ln(\tau)$ (first row in Fig. 9) decreases first if only one measurement is made, suggesting that the selected best channel is always sensitive to cloud optical thickness and uncertainty of $\ln(\tau)$ can be greatly reduced with only one measurement. Uncertainties of r_{eff} and h begin to reduce if two or more measurements are made. Furthermore, uncertainty of $\ln(\tau)$ decreases slowly for optically thin cloud (black curve in panel a) and rapidly for high cloud (red curve in panel c). Uncertainty of r_{eff} does not change significantly for optically thin and thick clouds (black and red curves in panel d), suggesting that IR observations have difficulty in retrieving r_{eff} if τ is too small or too large. For a transparent cirrus cloud, uncertainty of r_{eff} decreases rapidly if r_{eff} is small, as shown with the black curve in panel e. The third row of Fig. 9 shows the uncertainties of h for different clouds. The maximum decrease rate of h uncertainty occurs when the cloud is opaque ($\tau > 5$). On the contrary, it is difficult to retrieve h with existing IR observations when τ is small. In this case, multiple solutions (or a wide range of solution space) may exist and the retrieval algorithm cannot give further details of cloud-top height information. It should be pointed out that another possible situation is when h is above the tropopause, where the ambient air temperature may not change significantly with height, and the retrieval algorithm cannot find an accurate h since in the Jacobian matrix $\mathbf{K}_{F/X}$ approaches 0. As shown in Eq. (8), zero sensitivity does not improve \mathbf{S}_p even if \mathbf{S}_y is small.

Although using less than 5 measurements can efficiently reduce retrieval uncertainty to the minimum as mentioned, the needed combination of IR measurements varies significantly

with cloud properties and the background atmosphere. Therefore, we suggest using all of the bands to maximize the overall information content if computational efficiency is acceptable.

6. Conclusion

We document an OE-based retrieval algorithm we developed to infer three key ice cloud properties (τ , r_{eff} , and h) from MODIS measurements and analyze the corresponding uncertainties. The retrieval algorithm is based on an accurate and computationally efficient forward model that simulates TOA radiances in the infrared spectral region. This model employs the CRTM clear-sky module, which mitigates simulation biases due to the shifted SRFs of several Aqua MODIS IR bands. Absorption, emission, and scattering of ice cloud particles are considered using pre-computed LUTs in the forward model, as discussed in Section 2 and previous studies. Meanwhile, the present model also simulates critical Jacobian matrices, such as $\mathbf{K}_{F/X}$, \mathbf{K}_{F/τ_g} , $\mathbf{K}_{F/B(T)}$, and $\mathbf{K}_{F/C}$, facilitating a more comprehensive and computationally efficient IR ice cloud retrieval framework for research and operational applications, which went beyond previous studies by providing quantitative estimation of uncertainties from the atmospheric state, surface, and ice particles.

Four types of uncertainties are considered in the retrieval algorithm: the measurement, fast RT model, the ancillary data fields (e.g., atmospheric state), and the assumed ice crystal habit uncertainties. Except for the measurement uncertainty, the other uncertainties are influenced by cloud properties and clear-sky conditions such as the atmospheric state and surface emission. Inter-comparisons between the four types of uncertainties show that: (1) ancillary parameter uncertainties are dominant; (2) the maximum impact of ancillary parameter uncertainties occur in absorption bands; (3) in window bands, the impact of ice particle habit

uncertainty becomes more important and is comparable to the impact of ancillary parameter uncertainties; (4) uncertainties from observations and the fast RT model simulations are relatively small compared with ice particle and ancillary uncertainties.

Information content analysis is conducted based on the aforementioned model and error analysis. Three major conclusions can be gained from the study, which agree well with previous studies [e.g., *Cooper et al.*, 2006]. First, retrieval of the three ice cloud properties using IR observations has the best performance when the cloud is moderately optically thick ($\tau \sim 1-5$), located at a high altitude, and consists of relatively small particles. Under these conditions, the overall information content using the best 5 observations (IR channels) can exceed 10 bits. Overall information content decreases if the cloud is optically thin ($\tau \sim 0.1$) because of the large uncertainty associated with h retrieval, and if the cloud consists of large particles ($r_{eff} > 50 \mu\text{m}$) or is optically thick ($\tau \sim 10$) because of the large uncertainty associated with r_{eff} retrieval. Second, the importance of each band varies with cloud properties. For example, the thermal IR window bands are more important for optically thin cirrus cloud consisting of relatively small particles, whereas the mid-wave IR window band (e.g., $3.8 \mu\text{m}$) is more important when the cloud is optically thick. Third, the overall information content increases with the increase of measurements. However, H increases slowly when more than three bands are involved, suggesting that the MODIS IR bands are not independent. Generally, it is difficult to select a fixed combination of several bands for ice cloud retrieval without the knowledge of cloud properties. Therefore, we suggest using all of the bands to maximize the overall information content if computational efficiency is acceptable.

Acknowledgement

497 The authors are grateful for support from the NASA Radiation Sciences Program. Chenxi Wang
498 would like to thank Drs. Quanhua Liu, David Groff, and Eva E. Borbas for their help on CRTM.
499 The computations in this study were performed on UMBC High Performance Computing
500 Facility (HPCF). The facility is supported by the U.S. National Science Foundation through the
501 MRI program (grant nos. CNS-0821258 and CNS-1228778) and the SCREMS program (grant
502 no. DMS 0821311), with additional substantial support from UMBC.

References:

- Baran, A. J. (2012), From the single-scattering properties of ice crystals to climate prediction: A way forward, *Atmos. Res.*, *112*, 45-69.
- Baum, B. A., P. Yang, A. J. Heymsfield, C. G. Schmitt, Y. Xie, A. Bansemer, Y. X. Hu, and Z. Zhang (2011), Improvements to shortwave bulk scattering and absorption models for the remote sensing of ice clouds, *J. Appl. Meteor. Climatol.*, *50*, 1037-1056.
- Baum, B. A., W. P. Menzel, R. A. Frey, D. Tobin, R. E. Holz, S. A. Ackerman, A. K. Heidinger, and P. Yang (2012), MODIS cloud top property refinements for Collection 6, *J. Appl. Meteor. Climatol.*, *51*, 1145-1163.
- Baum, B. A., P. Yang, A. J. Heymsfield, A. Bansemer, A. Merrelli, C. Schmitt, and C. Wang (2014), Ice cloud bulk single-scattering property models with the full phase matrix at wavelengths from 0.2 to 100 μm , *J. Quant. Spectrosc. Radiat. Transfer*, *146*, 123-139, doi:10.1016/j.jqsrt.2014.02.029.
- Chepfer, H., G. Brogniez, and Y. Fouquart (1998), Cirrus clouds' microphysical properties deduced from POLDER observations, *J. Quant. Spectrosc. Radiat. Transfer*, *60*, 375-390.
- Coddington, O., P. Pilewskie, and T. Vukicevic (2012), The Shannon information content of hyperspectral shortwave cloud albedo measurements: Quantification and practical applications, *J. Geophys. Res.*, *117*, D04205, doi:10.1029/2011JD016771.
- Cole, B. H., P. Yang, B. A. Baum, J. Riedi, L. C.-Labonnote, F. Thieuleux, and S. Platnick (2013), Comparison of PARASOL observations with polarized reflectances simulated using different ice habit mixtures, *J. Appl. Meteor. Climatol.*, *52*, 186-196.

524 Cooper, S. J., T. S. L'Ecuyer, and G. L. Stephens (2003), The impact of explicit cloud boundary
525 information on ice cloud microphysical property retrievals from infrared radiances, *J.*
526 *Geophys. Res.*, *108*, 4107, doi:10.1029/2002JD002611.

527 Cooper, S. J., T. S. L'Ecuyer, P. Gabriel, A. J. Baran, and G. L. Stephens (2006), Objective
528 assessment of the information content of visible and infrared radiance measurements for
529 cloud microphysical property retrievals over the global oceans. Part II: ice cloud, *J. Appl.*
530 *Meteor. Climatol.*, *45*, 42-62.

531 Cornet, C., J. C. Buriez, J. Riedi, H. Isaka, and B. Guillemet (2005), Case study of
532 inhomogeneous cloud parameter retrieval from MODIS data, *Geophys. Res. Lett.*, *32*,
533 L13807.

534 Cox, C., and W. Munk (1954), Measurement of the roughness of the sea surface from
535 photographs of the sun's glitter, *J. Opt. Amer. Soc.*, *44*, 838-850.

536 de Reus, et al. (2009), Evidence for ice particles in the tropical stratosphere from in-situ
537 measurements, *Atmos. Chem. Phys.*, *9*, 6775-6792.

538 Fauchez, T., C. Cornet, F. Szczap, P. Dubuisson, and T. Rosambert, Impact of cirrus clouds
539 heterogeneities on top-of-atmosphere thermal infrared radiation (2014), *Atmos. Chem.*
540 *Phys.*, *14*, 5599-5615.

541 Fauchez, T., P. Dubuisson, C. Cornet, F. Szczap, A. Garnier, J. Pelon, and K. Meyer, Impacts of
542 cloud heterogeneities on cirrus optical properties retrieved from space-based thermal
543 infrared radiometry (2015), *Atmos. Meas. Tech.*, *8*, 633-647.

544 Garnier, A., J. Pelon, P. Dubuisson, M. Faivre, O. Chomette, N. Pascal, and D. P. Kratz (2012),
545 Retrieval of cloud properties using CALIPSO Imaging Infrared Radiometer. Part I:
546 Effective emissivity and optical depth, *J. Appl. Meteor. Climatol.*, *51*, 1407-1425.

547 Garnier, A., et al. (2013), Retrieval of cloud properties using CALIPSO Imaging Infrared
 548 Radiometer. Part II: Effective diameter and ice water path, *J. Appl. Meteor. Climatol.*, 52,
 549 2582-2599.

550 Han, Y., et al. (2006), Community radiative transfer model (CRTM) – version 1, *Technical*
 551 *report 122*, NOAA NESDIS.

552 Heidinger, A. K., and M. J. Pavolonis (2009), Gazing at cirrus clouds for 25 years through a split
 553 window. Part I: Methodology, *J. Appl. Meteor. Climatol.*, 48, 1100-1116,
 554 doi:10.1175/2008JAMC1882.1.

555 Heidinger, A. K., Y. Li, B. A. Baum, R. E. Holz, S. Platnick, and P. Yang (2015), Retrieval of
 556 cirrus cloud optical depth under day and night conditions from MODIS collection 6 cloud
 557 property data, *Remote. Sens.*, 7, 7257-7271, doi:10.3390/rs70607257.

558 Heymsfield, A. J., A. Bansemer, P. R. Field, S. L. Durden, J. L. Stith, J. E. Dye, W. Hall, and C.
 559 A. Grainger (2002), Observations and parameterizations of particle size distributions in
 560 deep tropical cirrus and stratiform precipitation clouds: Results from in situ observations
 561 in TRMM field campaigns, *J. Atmos. Sci.*, 59, 3457-3491.

562 Holz, R. E., et al. (2015), Resolving cirrus optical thickness biases between CALIOP and
 563 MODIS using infrared retrievals, *submitted to Atmos. Chem. Phys.*

564 Huang, H.-L., P. Yang, H. Wei, B. A. Baum, Y. X. Hu, P. Antonelli, and S. A. Ackerman
 565 (2004), Retrieval of ice cloud properties from high spectral resolution infrared
 566 observations, *IEEE Trans. Geosci. Remote Sens.*, 42, 842-853.

567 Inoue, T. (1985), On the temperature and effective emissivity determination of semitransparent
 568 cirrus clouds by bi-spectral measurements in the 10 μm window region, *J. Meteor. Soc.*
 569 *Japan*, 63, 88-89.

570 Iwabuchi, H, and T. Hayasaka (2002), Effects of cloud horizontal inhomogeneity on the optical
571 thickness retrieved from moderate-resolution satellite data, *J. Atmos. Sci.*, *59*, 2227-2242.

572 Iwabuchi, H., S. Yamada, S. Katagiri, P. Yang, and H. Okamoto (2014), Radiative and
573 microphysical properties of cirrus cloud inferred from infrared measurements made by
574 the Moderate Resolution Imaging Spectroradiometer (MODIS). Part I: Retrieval method,
575 *J. Appl. Meteor. Climatol.*, *53*, 1297-1316, doi:10.1175/JAMC-D-13-0215.1.

576 Kahn, B. H., C. K. Liang, A. Eldering, A. Gettelman, Q. Yue, and K. N. Liou (2008), Tropical
577 thin cirrus and relative humidity observed by the Atmospheric Infrared Sounder. *Atmos.*
578 *Chem. Phys.*, *8*, 1501-1518.

579 Kahn, B. H., et al. (2014), The Atmospheric Infrared Sounder version 6 cloud products, *Atmos.*
580 *Chem. Phys.*, *14*, 399-426.

581 King, M. D., Y. J. Kaufman, W. P. Menzel, and D. Tanre (1992), Remote sensing of cloud,
582 aerosol, and water vapor properties from the Moderate Resolution Imaging Spectrometer
583 (MODIS), *IEEE Trans. Geosci. Remote Sens.*, *30*, 2-27.

584 Kratz, D. P. (1995), The correlated k -distribution technique as applied to the AVHRR channels,
585 *J. Quant. Spectrosc. Radiat. Transfer*, *53*, 501-517.

586 L'Ecuyer, T. S., P. Gabriel, K. Leesman, S. J. Cooper, and G. L. Stephens (2006), Objective
587 assessment of the information content of visible and infrared radiance measurements for
588 cloud microphysical property retrievals over the global oceans. Part I: Liquid clouds, *J.*
589 *Appl. Meteor. Climatol.*, *45*, 20-41.

590 Levenberg, K. (1944), A method for the solution of certain non-linear problems in least squares,
591 *Quart. Appl. Math.*, *2*, 164-168.

592 Li, J.-L. F., et al. (2012), An observationally based evaluation of cloud ice water in CMIP3 and
 593 CMIP5 GCMs and contemporary reanalyses using contemporary satellite data, *J.*
 594 *Geophys. Res.*, *117*, D16105, doi:10.1029/2012JD017640.

595 Liou, K. N. (1986), Influence of cirrus clouds on weather and climate processes: A global
 596 perspective, *Mon. Wea. Rev.*, *114*, 1167-1199.

597 Liu, C., R. Panetta, and P. Yang (2013): The effects of surface roughness on the scattering
 598 properties of hexagonal columns with sizes from the Rayleigh to the geometric optics
 599 regimes, *J. Quant. Spectrosc. Radiat. Transfer*, *129*, 169-185,
 600 doi:10.1016/j.jqsrt.2013.06.011.

601 Liu, Q., and F. Weng (2006), Advanced doubling-adding method for radiative transfer in
 602 planetary atmosphere, *J. Atmos. Sci.*, *63*, 3459-3465.

603 Liu, Q., and S. Boukabara (2014), Community Radiative Transfer Model (CRTM) applications
 604 in supporting the Suomi National Polar-orbiting Partnership (SNPP) mission validation
 605 and verification, *Remote Sens. Environ.*, *140*, 744-754.

606 Marquardt, D. (1963), An algorithm for least-squares estimation of nonlinear parameters, *SIAM*
 607 *J. Appl. Math.*, *11*(2), 431-441.

608 Marshak, A., J. V. Martins, V. Zubko, and Y. J. Kaufman (2006), What does reflection from
 609 cloud sides tell us about vertical distribution of cloud droplet sizes, *Atmos. Chem. Phys.*,
 610 *6*, 5295-5305.

611 Menzel, W. P., W. L. Smith, and T. R. Stewart (1983), Improved cloud motion wind vector and
 612 altitude assignment using VAS, *J. Appl. Meteorol.*, *22*, 377-384.

613 Meyer, K. G., P. Yang, and B. -C. Gao (2004), Optical thickness of tropical cirrus clouds derived
614 from the MODIS 0.66- and 1.375- μm channels, *IEEE Trans. Geosci. Remote Sens.*, *42*,
615 833-841, doi:10.1109/TGRS.2003.818939.

616 Meyer, K. G., and S. Platnick (2010), Utilizing the MODIS 1.38 μm channel for cirrus cloud
617 optical thickness retrievals: Algorithm and retrieval uncertainties, *J. Geophys. Res.*, *115*,
618 D24209, doi:10.1029/2010JD014872.

619 Minnis, P., P. W. Heck, and D. F. Young (1993), Inference of cirrus cloud properties from
620 satellite-observed visible and infrared radiances. Part II: verification of theoretical
621 radiative properties, *J. Atmos. Sci.*, *50*, 1305-1322.

622 Minnis, P., D. P. Garber, D. F. Yong, R. F. Arduini, and Y. Takano (1998), Parameterizations of
623 reflectance and Effective emittance for satellite remote sensing of cloud properties, *J.*
624 *Atmos. Sci.*, *55*, 3313-3339.

625 Minnis, P., et al. (2011), CERES edition-2 cloud property retrievals using TRMM VIRS and
626 TERRA and AQUA MODIS data, part I: algorithms, *IEEE Trans. Geosci. Remote Sens.*,
627 *49*, 4374-4400.

628 Nakajima, T., and M. D. King (1990), Determination of the optical thickness and effective
629 particle radius of clouds from reflected solar radiance measurements. Part I: Theory, *J.*
630 *Atmos. Sci.*, *47*, 1878-1893.

631 Parol, F., J. C. Buriez, G. Brogniez, and Y. Fouquart (1991), Information content of AVHRR
632 channels 4 and 5 with respect to the effective radius of cirrus cloud particles, *J. Appl.*
633 *Meteorol.*, *30*, 973-984.

634 Pfalzgraff, W. C., R. M. Hulscher, and S. P. Neshyba (2010), Scanning electron microscopy and
 635 molecular dynamics of surfaces of growing and ablating hexagonal ice crystals, *Atmos.*
 636 *Chem. Phys.*, *10*, 2927-2935.

637 Poulsen, C. A., et al. (2012), Cloud retrievals from satellite data optimal estimation: Evaluation
 638 and application to ATSR, *Atmos. Meas. Tech.*, *5*, 1889-1910, doi:10.5194/amt-5-1889-
 639 2012.

640 Platnick, S., M. King, S. Ackerman, W. Menzel, B. Baum, J. Riedi, and R. Frey (2003), The
 641 MODIS cloud products: Algorithms and examples from Terra, *IEEE Trans. Geosci.*
 642 *Remote Sens.*, *41*, 459-473, doi:10.1109/TGRS.2002.808301.

643 Rodgers, C. D. (2000), *Inverse methods for atmospheric sounding*. World Scientific, 238 pp.

644 Shannon, C. E., and W. Weaver (1949), *The Mathematical Theory of Communication*. Paperback
 645 edition, University of Illinois Press, Urbana.

646 Smith, W. L., and C. M. R. Platt (1978), Intercomparison of radiosonde, ground based laser, and
 647 satellite deduced cloud heights, *J. Appl. Meteor.*, *17*, 1796-1802.

648 Sohn, B. J., B. R. Kim, and S. S. Lee (2010), Possible shift of spectral response function of the
 649 MODIS 6.8 μm water vapor channel causing a cold bias of 2-3 K, *Atmos. Meas. Tech.*, *3*,
 650 1667-1672, doi:10.5194/amt-3-1667-2010.

651 Sourdeval, O., L. C.-Labonnote, G. Brogniez, O. Jourdan, J. Pelon, and A. Garnier (2013), A
 652 variational approach for retrieving ice cloud properties from infrared measurements:
 653 application in the context of two IIR validation campaigns, *Atmos. Chem. Phys.*, *13*,
 654 8229-8244, doi:10.5194/acp-13-8229-2013.

655 Sourdeval, O., L. C.-Labonnote, A. J. Baran, and G. Brogniez (2015), A methodology for
656 simultaneous retrieval of ice and liquid water cloud properties. Part I: Information
657 content and case study, *Q. J. Roy. Meteor. Soc.*, *141*, 870-882.

658 Stamnes, K., S. Tsay, W. Wiscombe, and K. Jayaweera (1988), Numerically stable algorithm for
659 discrete-ordinate-method radiative-transfer in multiple-scattering and emitting layered
660 media. *Appl. Opt.*, *27*, 2502-2509.

661 Tobin, D. C., H. E. Revercomb, C. C. Moeller, and T. S. Pagano (2006), Use of Atmospheric
662 Infrared Sounder high-spectral resolution spectra to assess the calibration of Moderate
663 resolution Imaging Spectroradiometer on EOS Aqua, *J. Geophys. Res.*, *111*, D09S05,
664 doi:10.1029/2005JD006095.

665 Waliser, D., et al. (2009), Cloud ice: A climate model challenge with signs and expectations of
666 progress, *J. Geophys. Res.*, *114*, D00A21, doi:10.1029/2008JD010015.

667 Walther, A., and A. K. Heidinger (2012), Implementation of the daytime cloud optical and
668 microphysical properties algorithm (DCOMP) in PATMOS-x, *J. Appl. Meteor. Climatol.*,
669 *51*, 1371-1390, doi:10.1175/JAMC-D-11-0108.1.

670 Wang, C., P. Yang, B. A. Baum, S. Platnick, A. K. Heidinger, Y. X. Hu, and R. E. Holz (2011),
671 Retrieval of ice cloud optical thickness and effective particle size using a fast infrared
672 radiative transfer model, *J. Appl. Meteor. Climatol.*, *50*, 2283-2297.

673 Wang, C., P. Yang, S. Platnick, A. K. Heidinger, B. A. Baum, T. Greenwald, Z. Zhang, and R. E.
674 Holz (2013), Retrieval of ice cloud properties from AIRS and MODIS observations based
675 on a fast high-spectral-resolution radiative transfer model, *J. Appl. Meteor. Climatol.*, *52*,
676 710-726.

677 Wang, C., P. Yang, A. Dessler, B. A. Baum, and Y. X. Hu (2014), Estimation of the cirrus cloud
678 scattering phase function from satellite observations, *J. Quant. Spectrosc. Radiat.*
679 *Transfer*, 138, 36-49.

680 Watts, P. D., R. Bennartz, and F. Fell (2011), Retrieval of two-layer cloud properties from
681 multispectral observations using optimal estimation, *J. Geophys. Res.*, 116, D16203,
682 doi:10.1029/2011JD015883.

683 Wei, H., P. Yang, J. Li, B. A. Baum, H. L. Huang, S. Platnick, Y. X. Hu, and L. Strow (2004),
684 Retrieval of semitransparent ice cloud optical thickness from Atmospheric Infrared
685 Sounder (AIRS) measurements, *IEEE Trans. Geosci. Remote Sens.*, 42, 2254-2266.

686 Wylie, D. P., and W. P. Menzel (1999), Eight years of high cloud statistics using HIRS, *J.*
687 *Climate*, 12, 170-184.

688 Xiong, X., B. N. Wenny, A. Wu, W. L. Barnes, and V. V. Solomonson (2009), Aqua MODIS
689 thermal emissive band on-orbit calibration, characterization, and performance, *IEEE*
690 *Trans. Geosci. Remote Sens.*, 47, 803-814.

691 Yang, P., and K. N. Liou (1998), Single-scattering properties of complex ice crystals in
692 terrestrial atmosphere, *Contr. Atmos. Phys.* 71(2), 223–248.

693 Yang, P., L. Zhang, G. Hong, S. L. Nasiri, B. A. Baum, H.-L. Huang, M. D. King, and S.
694 Platnick (2007), Differences between collection 004 and 005 MODIS ice cloud
695 optical/microphysical products and their impact on radiative forcing simulations, *IEEE*
696 *Trans. Geosci. Remote Sens.*, 45, 2886-2899.

697 Yang, P., L. Bi, B. Baum, K. Liou, G. Kattawar, M. Mishchenko, and B. Cole (2013), Spectrally
698 Consistent Scattering, Absorption, and Polarization Properties of Atmospheric Ice

699 Crystals at Wavelengths from 0.2 to 100 μm , *J. Atmos. Sci.*, 70, 330-347,
700 doi:10.1175/JAS-D-12-039.1.

701 Yang, P., K. N. Liou, L. Bi, C. Liu, B. Q. Yi, and B. A. Baum (2015), On the radiative properties
702 of ice clouds: Light scattering, remote sensing, and radiation parameterization, *Adv.*
703 *Atmos. Sci.*, 32, 32–63, doi:10.1007/s00376-014-0011-z.

704 Zhang, Z., P. Yang, G. W. Kattawar, J. Riedi, L. C-Labonnote, B. A. Baum, S. Platnick, and H.
705 L. Huang (2009), Influence of ice particle model on satellite ice cloud retrieval: lessons
706 learned from MODIS and POLDER cloud product comparison, *Atmos. Chem. Phys.*, 9,
707 7115-7129.

708

709

710

711 **Appendix A**

712 Derivation of Jacobian matrix $\mathbf{K}_{F/X}$:

713 A quasi-analytical method is used to derive $\mathbf{K}_{F/X}$ in the forward model. For a single-layer
 714 cloudy atmosphere, the downward radiance reaching the surface can be approximately expressed
 715 as

$$716 \quad I_{surf}^{\downarrow}(\mu) = B(T_{eff-c}^{\downarrow})\varepsilon_{-c}(\mu)\Gamma_{gl}(\mu) + I_{gl}^{\downarrow}(\mu), \quad (A1)$$

717 where μ is the cosine of radiance zenith angle (absolute value), ε_{-c} is cloud emissivity, T_{eff-c} is
 718 cloud effective temperature, Γ_{gl} is the transmittance of the lower atmosphere, and I_{gl}^{\downarrow} is
 719 downward clear-sky emission of the lower atmosphere. Here we assume the cloud internal
 720 temperature varies linearly with height. The effective temperature of the cloud layer is defined
 721 as:

$$722 \quad T_{eff-c}^{\uparrow\downarrow}(\mu, T_1, T_2) = B^{-1}\left(\frac{I_{-c}^{\uparrow\downarrow}(\mu)}{\varepsilon_{-c}(\mu)}\right), \quad (A2)$$

723 where B^{-1} is the inverse Planck function, T_1 and T_2 are the temperatures at the cloud upper and
 724 lower boundaries, and $I_{-c}^{\uparrow\downarrow}(\mu)$ is the outgoing upward (or downward) cloud emission at the cloud
 725 layer boundary. In order to derive $T_{eff-c}^{\uparrow\downarrow}$, we first calculate the outgoing cloud emission $I_{-c}^{\uparrow\downarrow}(\mu)$
 726 with reference boundary temperatures $T_{ref,1}$ (upper: 245 K) and $T_{ref,2}$ (lower: 250 K) by using the
 727 32-Stream DISORT, and follow by deriving an effective temperature $T_{eff-c,ref}^{\uparrow\downarrow}$ of the reference
 728 layer with Eq. (A2). Finally, $T_{eff-c}^{\uparrow\downarrow}(\mu, T_1, T_2)$ for a cloud layer with arbitrary boundary
 729 temperatures T_1 and T_2 can be derived with a linear interpolation:

$$730 \quad T_{eff-c}^{\uparrow\downarrow} = T_1 + (T_2 - T_1) \times \frac{T_{eff-c,ref}^{\uparrow\downarrow} - T_{ref,1}}{T_{ref,2} - T_{ref,1}}. \quad (A3)$$

Cloud effective temperature is discussed in detail in *Wang et al.* [2013; 2014].

The upward radiance at the surface is

$$I_{surf}^{\uparrow} = B(T_{-s})\varepsilon_{-s} + \frac{2(1-\varepsilon_{-s})}{\pi} \int_0^1 I_{surf}^{\downarrow}(\mu) \mu d\mu, \quad (A4)$$

where T_{-s} is the surface temperature, ε_{-s} is surface emissivity, and I_{surf}^{\downarrow} is the downward radiance derived with Eq. (A1). The upward radiance at cloud bottom I_{cb}^{\uparrow} , omitting multiple reflections between cloud and the surface, is given by

$$I_{cb}^{\uparrow}(\mu) = I_{surf}^{\uparrow} \Gamma_{gl}(\mu) + I_{gl}^{\uparrow}(\mu), \quad (A5)$$

where I_{gl}^{\uparrow} is the upward clear-sky emission of the lower atmosphere. The upward radiance at cloud-top I_{ct}^{\uparrow} consists of 1) diffusely and directly transmitted upward radiances from cloud bottom, 2) cloud upward emission, and 3) reflected downward radiance from the upper atmosphere:

$$I_{ct}^{\uparrow}(\mu_v) = 2 \int_0^1 [I_{cb}^{\uparrow}(\mu) \Gamma_{-c}(\mu, \mu_v) + I_{gh}^{\downarrow}(\mu) R_{-c}(\mu, \mu_v)] \mu d\mu + I_{cb}^{\uparrow}(\mu_v) e^{-\frac{\tau}{\mu_v}} + B(T_{eff-c}^{\uparrow}) \varepsilon_{-c}(\mu_v), \quad (A6)$$

where μ_{ϖ} is the cosine of the viewing zenith angle, I_{gh}^{\downarrow} is the upward clear-sky emission of the above cloud atmosphere, Γ_{-c} and R_{-c} are the azimuthal averaged bi-directional transmittance and reflectance of the cloud layer, respectively, and $e^{-\frac{\tau}{\mu_v}}$ is the cloud direct transmittance. The satellite observed radiance is

$$I_{TOA}^{\uparrow}(\mu_v) = I_{ct}^{\uparrow}(\mu_v) \Gamma_{gh}(\mu_v) + I_{gh}^{\uparrow}(\mu_v), \quad (A7)$$

where I_{gh}^{\uparrow} is the upward clear-sky emission of the upper atmosphere and Γ_{gh} is the transmittance of the upper atmosphere. Satellite observations can be simulated using Equations (A4)-(A7) with

known cloud and clear-sky optical properties. In these equations, ε_{-c} , Γ_{-c} , R_{-c} , and T_{eff-c} are functions of τ and r_{eff} . Additionally, I_{gl}^\uparrow , I_{gl}^\downarrow , I_{gh}^\uparrow , I_{gh}^\downarrow , Γ_{gl} , and Γ_{gh} are functions of the cloud-top height (h).

It is straightforward to derive the analytical expression of the Jacobian matrix $\mathbf{K}_{F/X}$ if the first order derivatives of the aforementioned variables with respect to τ , r_{eff} , and h are given. For example, the derivative of I_{surf}^\downarrow with respect to τ is given by

$$\frac{\partial I_{surf}^\downarrow}{\partial \tau} = \Gamma_{gl}(\mu) \left[B(T_{eff-c}^\downarrow) \frac{\partial \varepsilon_{-c}(\mu)}{\partial \tau} + \frac{\partial B}{\partial T} \frac{\partial T_{eff-c}^\downarrow}{\partial \tau} \varepsilon_{-c}(\mu) \right]. \quad (A8)$$

Similarly, derivatives of Equations (A4) ~ (A7) can be expressed as

$$\frac{\partial I_{surf}^\uparrow}{\partial \tau} = \frac{2(1-\varepsilon_{-s})}{\pi} \int_0^1 \frac{\partial I_{surf}^\downarrow(\mu)}{\partial \tau} \mu d\mu, \quad (A9)$$

$$\frac{\partial I_{cb}^\uparrow(\mu)}{\partial \tau} = \frac{\partial I_{surf}^\uparrow}{\partial \tau} \Gamma_{gl}(\mu), \quad (A10)$$

$$\begin{aligned} \frac{\partial I_{ct}^\uparrow(\mu_v)}{\partial \tau} = & 2 \int_0^1 \left[\frac{\partial I_{cb}^\uparrow(\mu)}{\partial \tau} \Gamma_{-c}(\mu, \mu_v) + I_{cb}^\uparrow(\mu) \frac{\partial \Gamma_{-c}(\mu, \mu_v)}{\partial \tau} \right] \mu d\mu + \\ & 2 \int_0^1 \left[I_{gh}^\downarrow(\mu) \frac{\partial R_{-c}(\mu, \mu_v)}{\partial \tau} \right] \mu d\mu + \frac{\partial I_{cb}^\uparrow(\mu_v)}{\partial \tau} e^{-\frac{\tau}{\mu_v}} + \\ & B(T_{eff-c}^\uparrow) \frac{\partial \varepsilon_{-c}(\mu_v)}{\partial \tau} + \frac{\partial B}{\partial t} \frac{\partial T_{eff-c}^\uparrow}{\partial \tau} \varepsilon_{-c}(\mu_v) \end{aligned} \quad (A11)$$

$$\frac{\partial I_{TOA}^\uparrow(\mu_v)}{\partial \tau} = \frac{\partial I_{ct}^\uparrow(\mu_v)}{\partial \tau} \Gamma_{gh}(\mu_v). \quad (A12)$$

Cloud scattering properties are pre-computed and stored in LUTs. The corresponding first order derivatives of the cloud optical properties shown in Equations (A8)-(A12) can be numerically derived using these LUTs.

767 Appendix B

768 Derivation of Jacobian matrices $\mathbf{K}_{\tau_g/\mathbf{T}}$, $\mathbf{K}_{\tau_g/\mathbf{w}}$, $\mathbf{K}_{\mathbf{F}/\tau_g}$, and $\mathbf{K}_{\mathbf{F}/\mathbf{B}(\mathbf{T})}$:

769 We derive the first two matrices numerically. Specifically, in one run of the forward RT
 770 model, we conduct three runs of the CRTM clear-sky module with unbiased (\mathbf{T} and \mathbf{w}) profiles,
 771 perturbed temperature profile ($\mathbf{T} + \Delta\mathbf{T}$ and \mathbf{w}), and perturbed water vapor concentration profile
 772 (\mathbf{T} and $\mathbf{w} + \Delta\mathbf{w}$), respectively. The temperature perturbation ΔT is a constant 0.1 K for all
 773 atmospheric layers, while $\Delta\mathbf{w}$ is assumed to be 5% of \mathbf{w} for all layers in practice. $\mathbf{K}_{\tau_g/\mathbf{T}}$ and
 774 $\mathbf{K}_{\tau_g/\mathbf{w}}$ are two $ml \times l$ matrices. Here m and l are the number of measurements and atmospheric
 775 layers, respectively. For example, $\mathbf{K}_{\tau_g/\mathbf{T}}$ can be expressed as a vector of m square sub-matrices:

$$776 \quad \mathbf{K}_{\tau_g/\mathbf{T}} = \left[\mathbf{K}_{\tau_g/\mathbf{T}}^1, \mathbf{K}_{\tau_g/\mathbf{T}}^2, \dots, \mathbf{K}_{\tau_g/\mathbf{T}}^i, \dots, \mathbf{K}_{\tau_g/\mathbf{T}}^m \right]^T, \quad (\text{B1})$$

777 where the superscript i indicates the i^{th} measurement, and the square sub-matrix $\mathbf{K}_{\tau_g/\mathbf{T}}^i$ ($l \times l$) is
 778 given by

$$779 \quad \mathbf{K}_{\tau_g/\mathbf{T}}^i = \begin{bmatrix} \frac{\partial \tau_{g1}^i}{\partial T_1} & \frac{\partial \tau_{g1}^i}{\partial T_2} & \dots & \frac{\partial \tau_{g1}^i}{\partial T_l} \\ \frac{\partial \tau_{g2}^i}{\partial T_1} & \frac{\partial \tau_{g2}^i}{\partial T_2} & \dots & \frac{\partial \tau_{g2}^i}{\partial T_l} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \tau_{gl}^i}{\partial T_1} & \frac{\partial \tau_{gl}^i}{\partial T_2} & \dots & \frac{\partial \tau_{gl}^i}{\partial T_l} \end{bmatrix} = \begin{bmatrix} \frac{\partial \tau_{g1}^i}{\partial T_1} & 0 & \dots & 0 \\ 0 & \frac{\partial \tau_{g2}^i}{\partial T_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{\partial \tau_{gl}^i}{\partial T_l} \end{bmatrix}, \quad (\text{B2})$$

780 where the subscript indicates the index of an atmospheric layer. For the j^{th} layer, $\frac{\partial \tau_{gj}^i}{\partial T_j}$ is given by

$$781 \quad \frac{\partial \tau_{gj}^i}{\partial T_j} \approx \frac{\tau_{gj}^i(T_j + \Delta T, w_j) - \tau_{gj}^i(T_j, w_j)}{\Delta T}. \quad (\text{B3})$$

782 A similar approach can be used to derive $\mathbf{K}_{\tau_g/\mathbf{w}}$.

783

784 The last two Jacobian matrices ($\mathbf{K}_{\mathbf{F}/\tau_g}$ and $\mathbf{K}_{\mathbf{F}/\mathbf{B}(\mathbf{T})}$) have m rows and $m \times l$ columns, both of
 785 which can be expressed as a block diagonal matrix:

$$786 \quad \mathbf{K}_{\mathbf{F}/\tau_g} = \begin{bmatrix} \mathbf{K}_{F^1/\tau_g^1} & \mathbf{K}_{F^1/\tau_g^2} & \cdots & \mathbf{K}_{F^1/\tau_g^m} \\ \mathbf{K}_{F^2/\tau_g^1} & \mathbf{K}_{F^2/\tau_g^2} & \cdots & \mathbf{K}_{F^2/\tau_g^m} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{K}_{F^m/\tau_g^1} & \mathbf{K}_{F^m/\tau_g^2} & \cdots & \mathbf{K}_{F^m/\tau_g^m} \end{bmatrix} = \begin{bmatrix} \mathbf{K}_{F^1/\tau_g^1} & 0 & \cdots & 0 \\ 0 & \mathbf{K}_{F^2/\tau_g^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{K}_{F^m/\tau_g^m} \end{bmatrix}. \quad (\text{B4})$$

787 Again, superscripts in Eq. (B4) indicate the indices of the measurements. Each sub-matrix in Eq.
 788 (B4) is a vector consisting of l elements:

$$789 \quad \mathbf{K}_{F^i/\tau_g^i} = \left[\frac{\partial F^i}{\partial \tau_{g1}^i}, \frac{\partial F^i}{\partial \tau_{g2}^i}, \dots, \frac{\partial F^i}{\partial \tau_{gj}^i}, \dots, \frac{\partial F^i}{\partial \tau_{gl}^i} \right]. \quad (\text{B5})$$

790 The j^{th} element $\frac{\partial F^i}{\partial \tau_{gj}^i}$ in vector $\mathbf{K}_{F^i/\tau_g^i}$ represents the sensitivity of the i^{th} measurement to the gas

791 optical thickness of the j^{th} layer. $\frac{\partial F^i}{\partial \tau_{gj}^i}$ can be analytically derived by differentiating Equations

792 (A4) ~ (A8) with respect to τ_{gj}^i . A similar approach can be used to derive $\mathbf{K}_{\mathbf{F}/\mathbf{B}(\mathbf{T})}$.

793

794

795

796 Appendix C

797 Derivation of Jacobian matrix $\mathbf{K}_{F/C}$:

798 Jacobian matrix $\mathbf{K}_{F/C}$ can be analytically derived by differentiating Eqs. (A1), and (A4) ~ (A8)
 799 with respect to the corresponding cloud scattering properties. For example, the derivative of I_{surf}^\downarrow
 800 with respect to ε_{-c} is given by

$$801 \quad \frac{\partial I_{surf}^\downarrow(\mu)}{\partial \varepsilon_{-c}(\mu)} = B(T_{eff_{-c}}^\downarrow) \Gamma_{gl}(\mu), \quad (C1)$$

802 Similarly, derivatives of Equations (A4) ~ (A8) with respect to $\varepsilon_{-c}(\mu_i)$ can be expressed as

$$803 \quad \frac{\partial I_{surf}^\uparrow}{\partial \varepsilon_{-c}(\mu_i)} = \frac{2(1 - \varepsilon_s)}{\pi} \int_0^1 \frac{\partial I_{surf}^\downarrow(\mu')}{\partial \varepsilon_{-c}(\mu')} \delta(\mu' - \mu_i) \mu' d\mu', \quad (C2)$$

$$804 \quad \frac{\partial I_{cb}^\uparrow(\mu)}{\partial \varepsilon_{-c}(\mu_i)} = \frac{\partial I_{surf}^\uparrow}{\partial \varepsilon_{-c}(\mu_i)} \Gamma_{gl}(\mu), \quad (C3)$$

$$805 \quad \frac{\partial I_{ct}^\uparrow(\mu_v)}{\partial \varepsilon_{-c}(\mu_i)} = 2 \int_0^1 \left[\frac{\partial I_{cb}^\uparrow(\mu)}{\partial \varepsilon_{-c}(\mu)} \Gamma_{-c}(\mu, \mu_v) \delta(\mu - \mu_i) \right] \mu d\mu + \frac{\partial I_{cb}^\uparrow(\mu_v)}{\partial \varepsilon_{-c}(\mu_i)} e^{-\frac{\tau}{\mu_v}} + B(T_{eff_{-c}}^\uparrow) \delta(\mu_i - \mu_v), \quad (C4)$$

$$806 \quad \frac{\partial I_{TOA}^\uparrow(\mu_v)}{\partial \varepsilon_{-c}(\mu_i)} = \frac{\partial I_{ct}^\uparrow(\mu_v)}{\partial \varepsilon_{-c}(\mu_i)} \Gamma_{gh}(\mu_v), \quad (C5)$$

807 where $\delta(\mu - \mu_i)$ is the Dirac Delta function, which is 0 if $\mu \neq \mu_i$, and infinity large if $\mu = \mu_i$.

808 Similar approaches are used to derive the derivative of TOA radiance with respect to other cloud
 809 parameters, such as reflectivity (R_{-c}) and transmissivity (Γ_{-c}).

810

811

812 **Table Captions**

813 Table 1. Geometries and ice cloud properties selected to calculate the cloud lookup tables

814 (LUTs). P43.

815

816 Table 1. Geometries and ice cloud properties selected to calculate the cloud lookup tables
817 (LUTs).

Ice Cloud LUTs	
16 Viewing/Incident angles (cosine)	0.99470, 0.97229, 0.93282, 0.87770, 0.80894, 0.72901, 0.64080, 0.54751, 0.45249, 0.35920, 0.27099, 0.19106, 0.12230, 0.06718, 0.02771, 0.00530.
40 Cloud Optical thickness	0.01, 0.03, 0.05, 0.10, 0.20, 0.30, 0.40, 0.50, 0.60, 0.70, 0.80, 0.90, 1.00, 1.20, 1.40, 1.60, 1.80, 2.00, 2.50, 3.00, 3.50, 4.00, 4.50, 5.00, 5.50, 6.00, 6.50, 7.00, 7.50, 8.00, 8.50, 9.00, 9.50, 10.0, 12.0, 15.0, 20.0, 25.0, 30.0, 50.0.
18 Cloud Effective Radius (μm)	5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90

818

819

820 **Figure Captions**

821 Figure 1. The forward model simulations of a single-layer cloudy (ice) atmosphere with $r_{eff} = 20$
822 μm , cloud-top height $h = 10$ km, surface temperature $T_{-s} = 294$ K, and surface emissivity
823 $\varepsilon_{-s} = 0.98$. Here TOA BT as a function of cloud optical thickness τ (panel a) is shown, as
824 is the first derivative of TOA BTs with respect to τ (panel b), r_{eff} (panel c), and h (panel
825 d). P46.

826 Figure 2. The three components (\mathbf{S}_{anc_s} , $\mathbf{S}_{anc_T_g}$, and \mathbf{S}_{anc_B}) of error covariance matrices \mathbf{S}_{anc} .
827 The surface temperature and emissivity uncertainties (δT_{-s} and $\delta \varepsilon_{-s}$) are 0.5K and 0.01,
828 respectively. The air temperature and water vapor concentration uncertainties (δT and
829 δw) are 1.0K and 15%, respectively, for all atmospheric layers. Note that the four types
830 of uncertainties δT_{-s} , $\delta \varepsilon_{-s}$, δT and δw are independent. Vertical correlations are
831 considered for both δT and δw . Calculations are for a single-layered ice cloud with $\tau = 1$,
832 $r_{eff} = 20\mu\text{m}$, and $h = 10\text{km}$. P47.

833 Figure 3. Same as Figure 2, but δT_{-s} and $\delta \varepsilon_{-s}$ are 2K and 0.03, respectively. The δT and δw are
834 2.0K and 25%, respectively, for all atmospheric layers. The other conditions are the same
835 as Figure 2. P48.

836 Figure 4. The three components (\mathbf{S}_{obs} , \mathbf{S}_{anc} , and \mathbf{S}_{habit}) of error covariance matrices \mathbf{S}_y . Note that
837 \mathbf{S}_{RT} is not shown in this figure since \mathbf{S}_{RT} is two orders of magnitude smaller than \mathbf{S}_{anc} . \mathbf{S}_{obs}
838 is generated using typical Aqua MODIS observations. Other conditions are the same as
839 Figure 2. P49.

840 Figure 5. Fractions of the four components of the diagonal elements in \mathbf{S}_y shown in Figure 4.
841 P50.

Figure 6. Uncertainties of τ , r_{eff} , and h due to measurement errors (red), uncertainties in ancillary datasets (yellow), ice cloud particle habit (light blue), and all four uncertainties combined (dark blue). Cloud and atmospheric conditions are the same as Figure 2. P51.

Figure 7. Information content H (in bits) for different band combinations and different cloud states. In each panel, the black curve indicates the H of a single band. The first best band with highest H can be found from the black curve. The blue curve indicates H of a combination of the current band and the first best band. Similarly, the second best band can be found from the blue curve. The green, yellow, and red curves indicate the H of three, four, and five bands. Corresponding first 5 best bands are given at the bottom of each panel. Ice cloud r_{eff} values are $15 \mu\text{m}$ for all panels. First row: high clouds with $h = 15 \text{ km}$; from left to right: τ increases from 0.1 to 5.0. Second row: moderately high clouds with $h = 10 \text{ km}$. P52.

Figure 8. Same as Figure 7, but ice cloud r_{eff} values are $50 \mu\text{m}$ for all panels. P53.

Figure 9. Uncertainty of each retrieval variable as a function of the number of measurements. First row: uncertainty of $\ln(\tau)$. Second row: uncertainty of $\ln(r_{eff})$. Third row: uncertainty of h . First column: fixed r_{eff} and h ($15 \mu\text{m}$ and 10 km , respectively), τ varies from 0.1 to 10.0. Second column: fixed τ and h (1 and 10 km , respectively), r_{eff} varies from 5 to $50 \mu\text{m}$. Third column: fixed τ and r_{eff} (1 and $15 \mu\text{m}$, respectively), h varies from 5 to 12.5 km . P54.

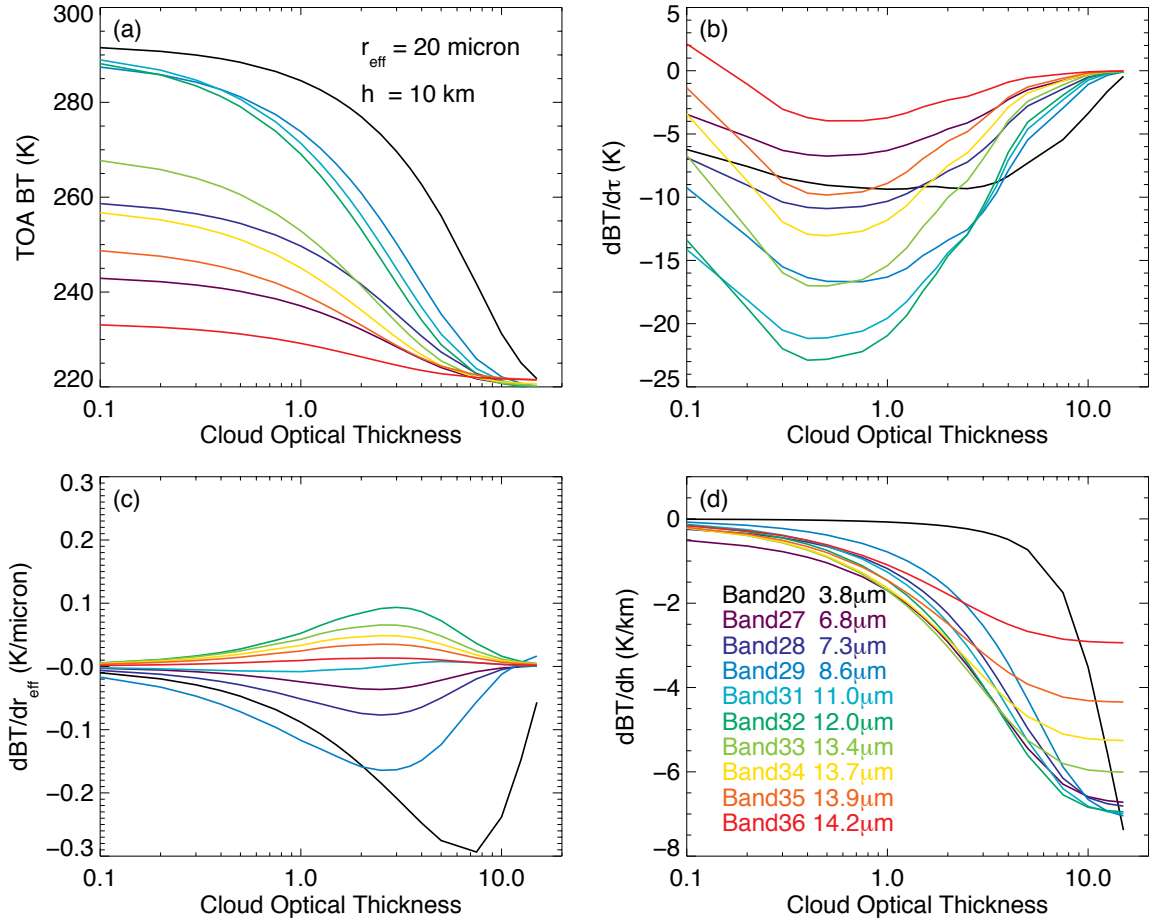


Figure 1. The forward model simulations of a single-layer cloudy (ice) atmosphere with $r_{\text{eff}} = 20 \mu\text{m}$, cloud-top height $h = 10 \text{ km}$, surface temperature $T_s = 294 \text{ K}$, and surface emissivity $\epsilon_s = 0.98$. Here TOA BT as a function of cloud optical thickness τ (panel a) is shown, as is the first derivative of TOA BTs with respect to τ (panel b), r_{eff} (panel c), and h (panel d).

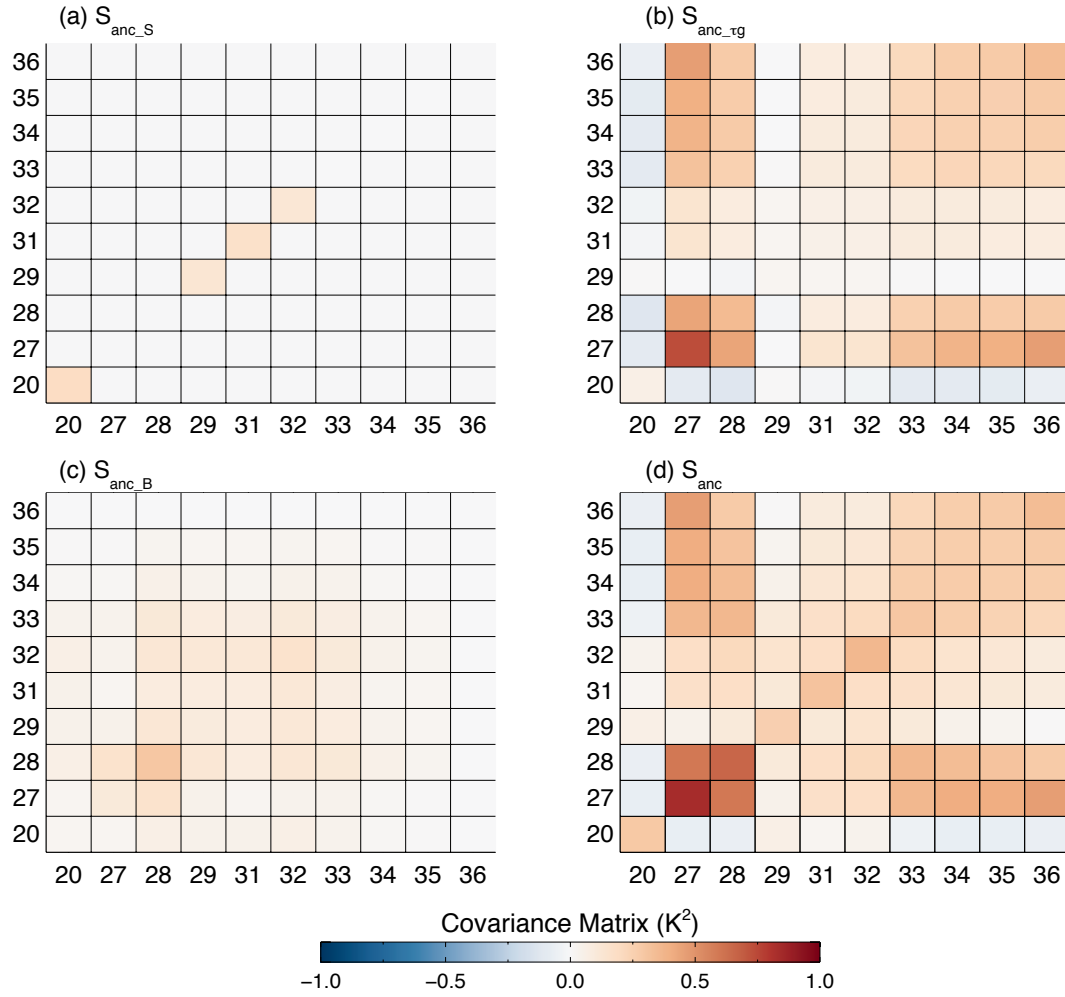


Figure 2. The three components (S_{anc_S} , S_{anc_rg} , and S_{anc_B}) of error covariance matrices S_{anc} . The surface temperature and emissivity uncertainties (δT_s and $\delta \epsilon_s$) are 0.5K and 0.01, respectively. The air temperature and water vapor concentration uncertainties (δT and δw) are 1.0K and 15%, respectively, for all atmospheric layers. Note that the four types of uncertainties δT_s , $\delta \epsilon_s$, δT and δw are independent. Vertical correlations are considered for both δT and δw . Calculations are for a single-layered ice cloud with $\tau = 1$, $r_{eff} = 20\mu m$, and $h = 10km$.

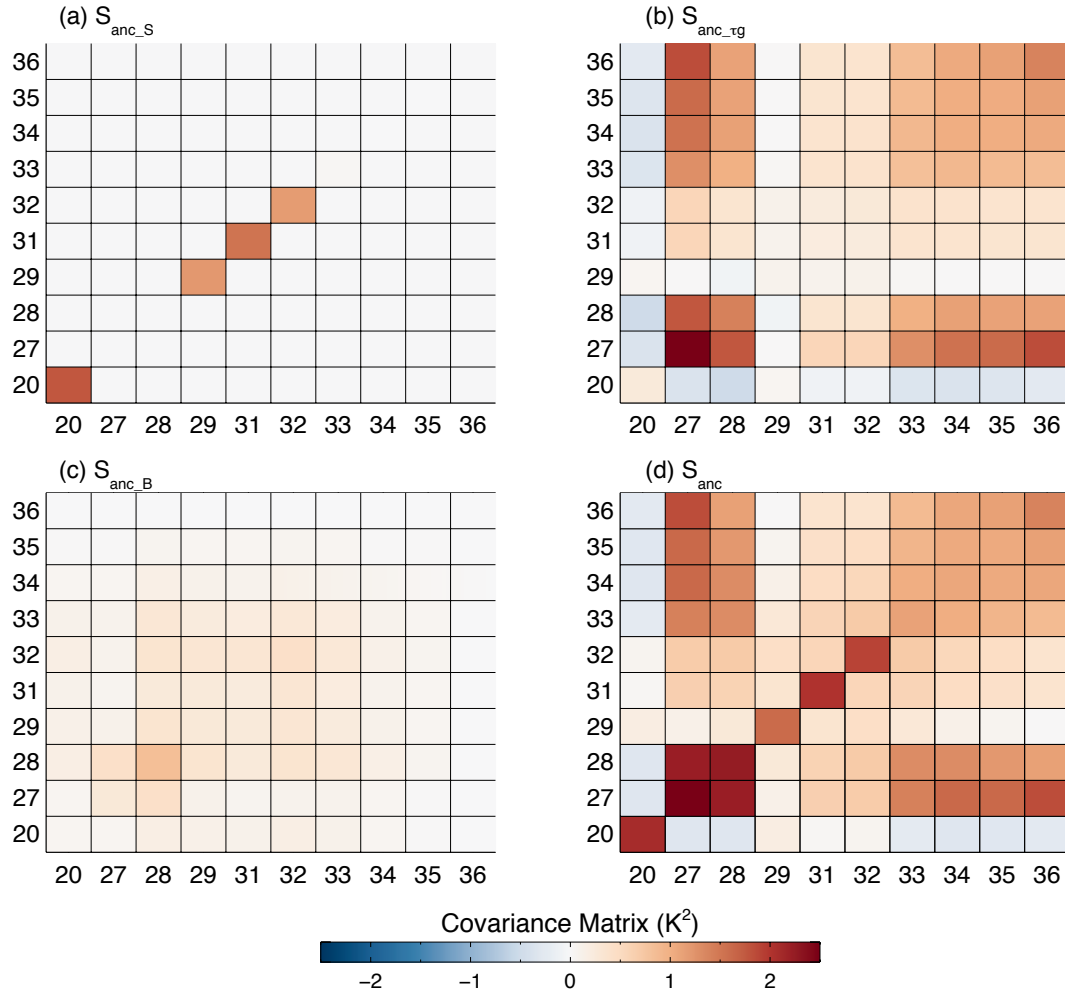


Figure 3. Same as Figure 2, but δT_{-s} and $\delta \varepsilon_{-s}$ are 2K and 0.03, respectively. The δT and δw are 2.0K and 25%, respectively, for all atmospheric layers. The other conditions are the same as Figure 2.

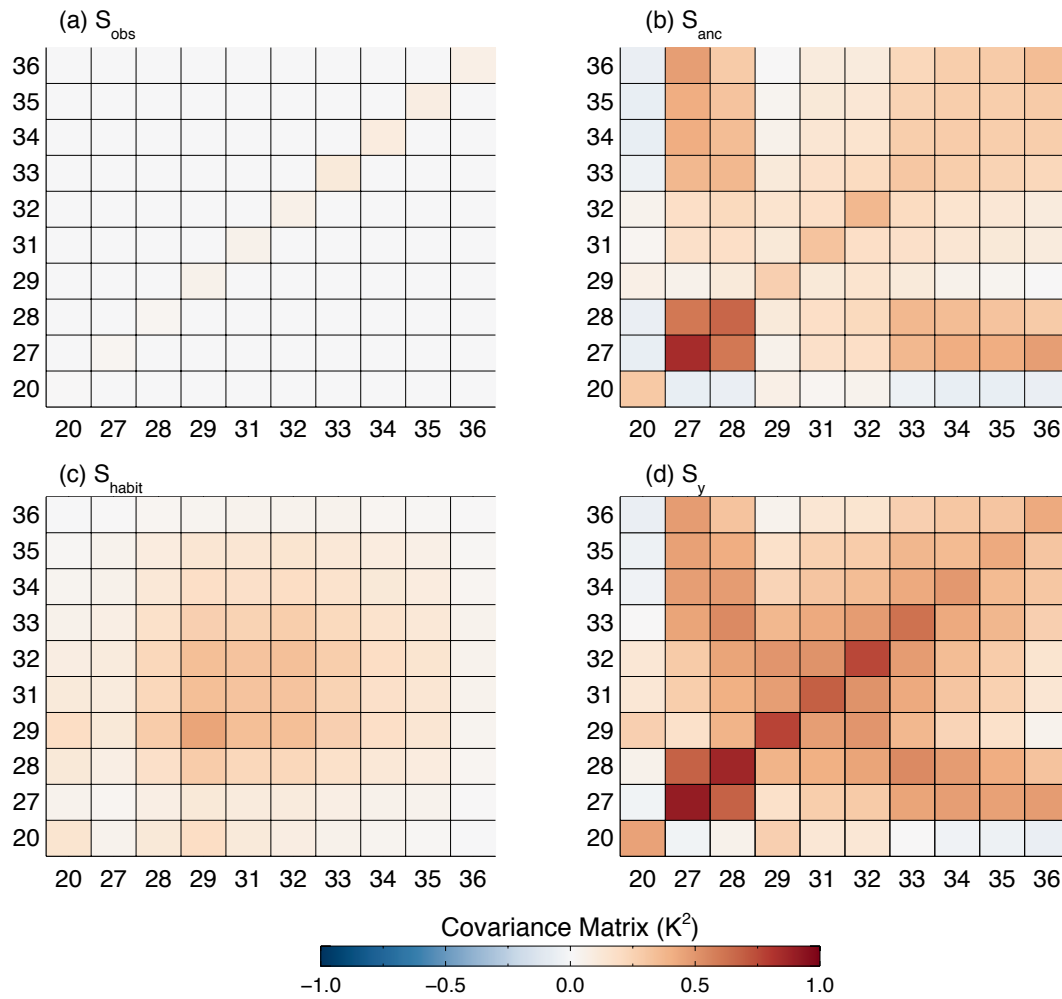


Figure 4. The three components (S_{obs} , S_{anc} , and S_{habit}) of error covariance matrices S_y . Note that S_{RT} is not shown in this figure since S_{RT} is two orders of magnitude smaller than S_{anc} . S_{obs} is generated using typical Aqua MODIS observations. Other conditions are the same as Figure 2.

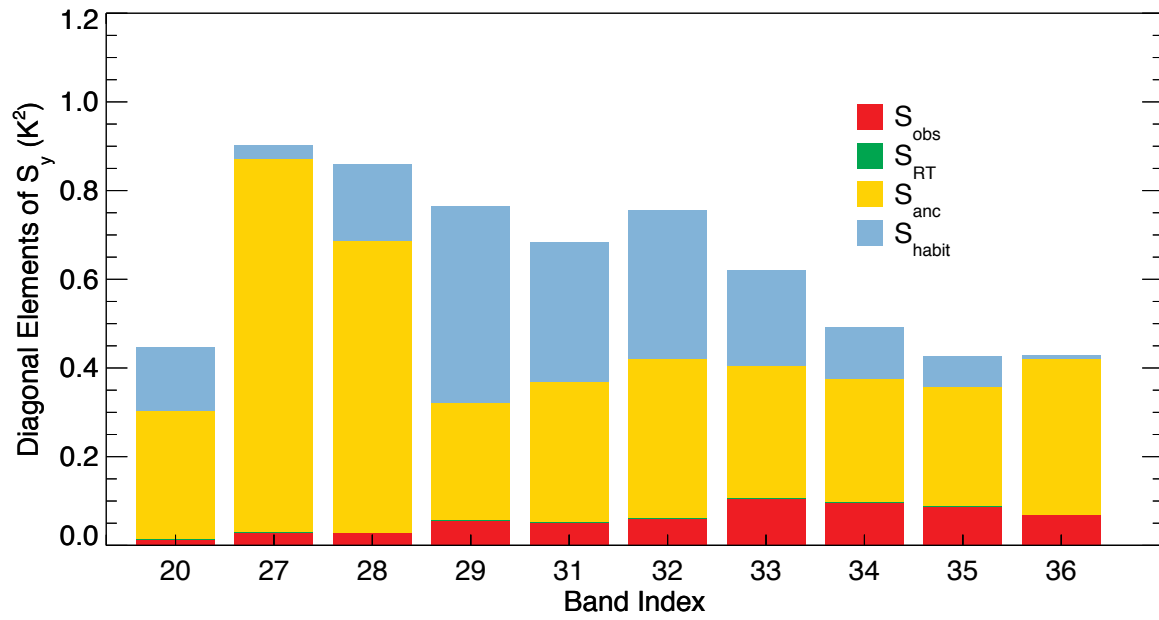


Figure 5. Fractions of the four components of the diagonal elements in S_y shown in Figure 4.

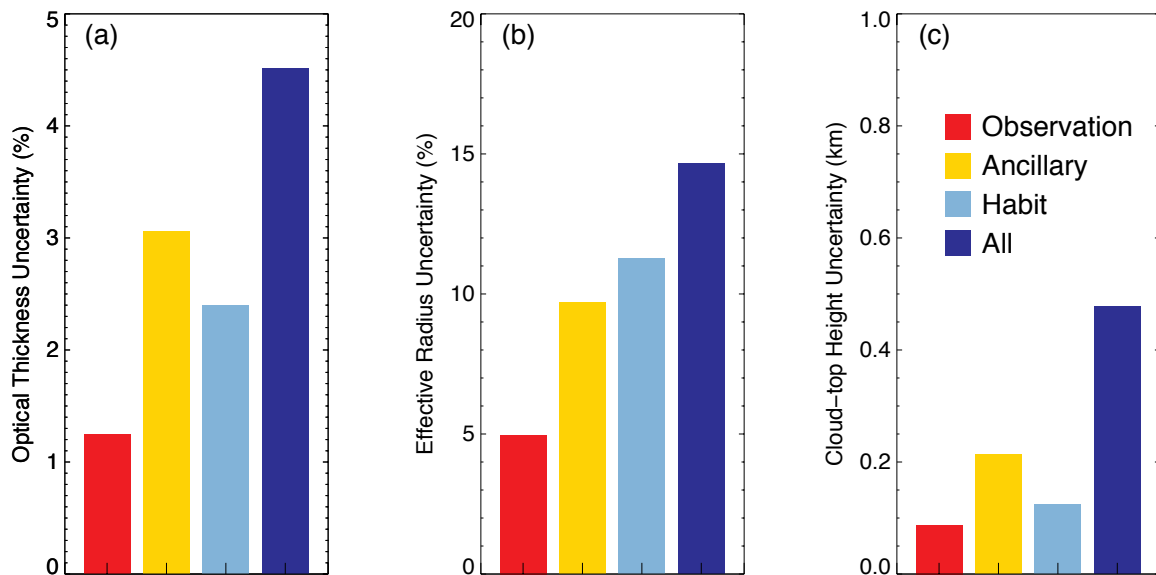


Figure 6. Uncertainties of τ , r_{eff} , and h due to measurement errors (red), uncertainties in ancillary datasets (yellow), ice cloud particle habit (light blue), and all four uncertainties combined (dark blue). Cloud and atmospheric conditions are the same as Figure 2. P47.

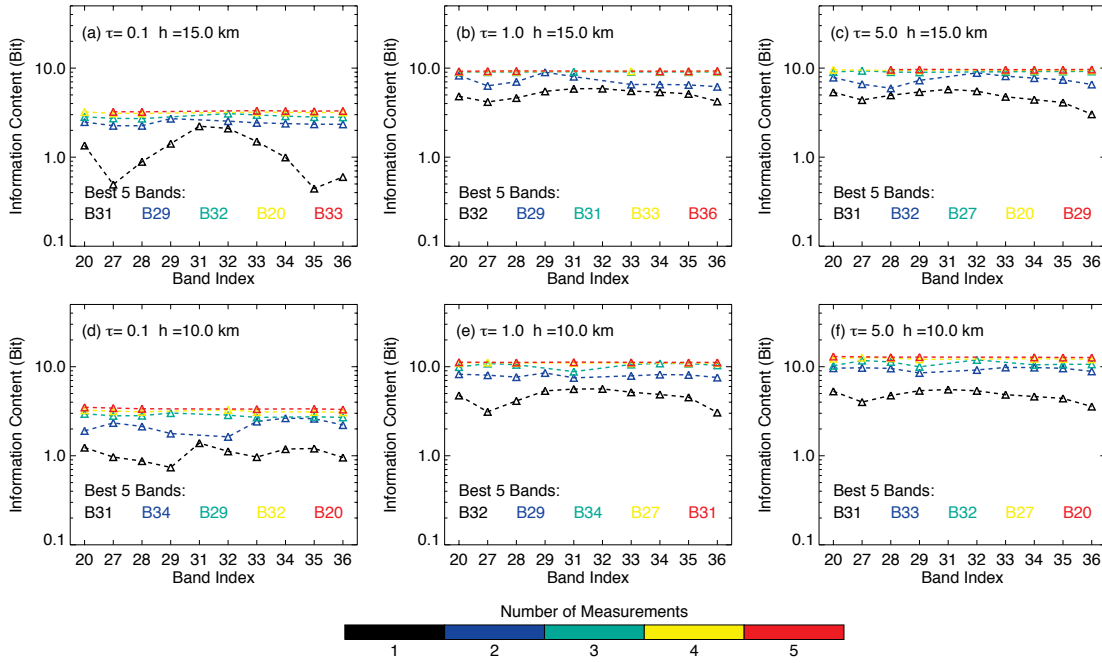


Figure 7. Information content H (in bits) for different band combinations and different cloud states. In each panel, the black curve indicates the H of a single band. The first best band with highest H can be found from the black curve. The blue curve indicates H of a combination of the current band and the first best band. Similarly, the second best band can be found from the blue curve. The green, yellow, and red curves indicate the H of three, four, and five bands. Corresponding first 5 best bands are given at the bottom of each panel. Ice cloud r_{eff} values are $15 \mu\text{m}$ for all panels. First row: high clouds with $h = 15$ km; from left to right: τ increases from 0.1 to 5.0. Second row: moderately high clouds with $h = 10$ km.

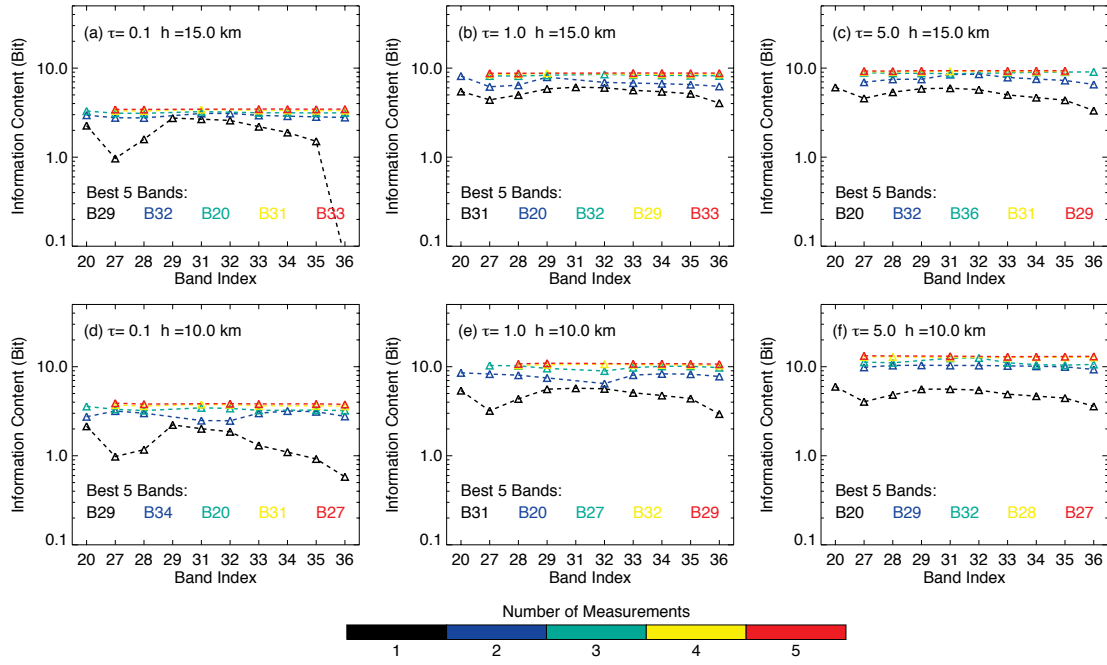


Figure 8. Same as Figure 7, but ice cloud r_{eff} values are $50 \mu\text{m}$ for all panels.

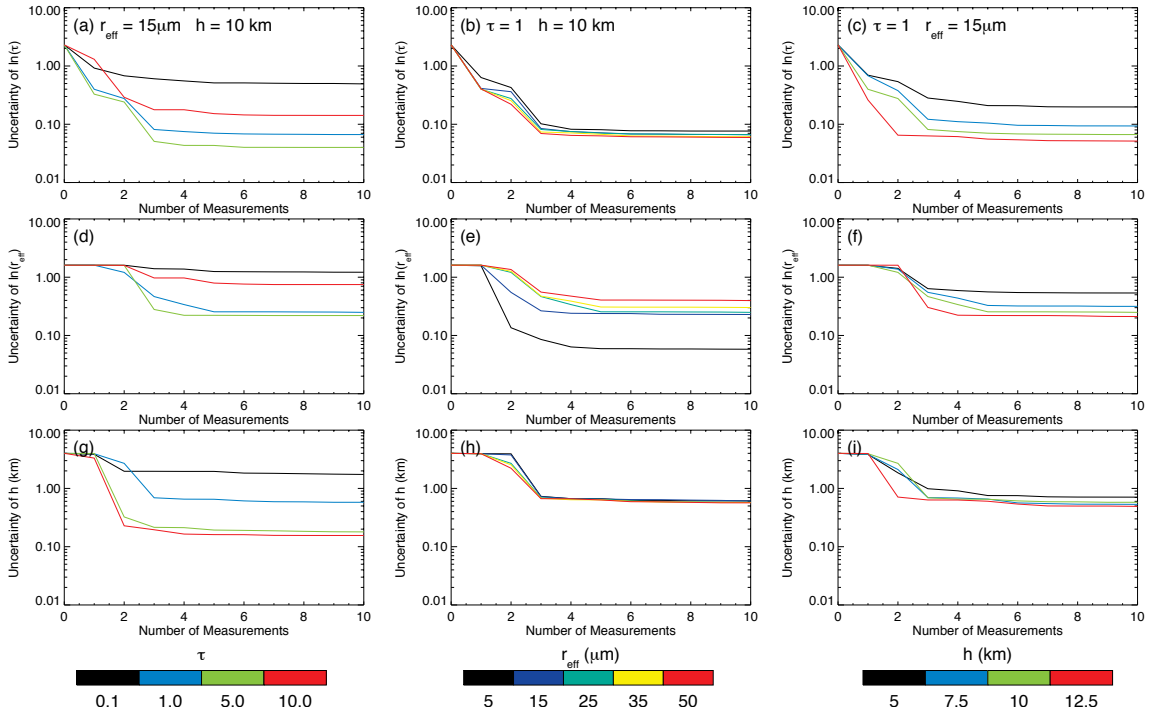


Figure 9. Uncertainty of each retrieval variable as a function of the number of measurements. First row: uncertainty of $\ln(\tau)$. Second row: uncertainty of $\ln(r_{eff})$. Third row: uncertainty of h . First column: fixed r_{eff} and h ($15\mu\text{m}$ and 10 km , respectively), τ varies from 0.1 to 10.0 . Second column: fixed τ and h (1 and 10 km , respectively), r_{eff} varies from 5 to $50\text{ }\mu\text{m}$. Third column: fixed τ and r_{eff} (1 and $15\mu\text{m}$, respectively), h varies from 5 to 12.5 km .